

05-25-00

PATENT
Attorney Docket No.: FORS-04323

A

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Box Patent Application

Assistant Commissioner For Patents
Washington, D.C. 20231

NEW APPLICATION TRANSMITTAL

Transmitted herewith for filing is the patent application of Wu-Po Ma, Victor I. Lyamichev, Michael W. Kaiser, Natalie E. Lyamicheva, Hatim Taysir Allawi, James J. Schaefer and Bruce P. Neri for **Improved Enzymes For The Detection Of Specific Nucleic Acid Sequences**.

CERTIFICATION UNDER 37 C.F.R. § 1.10

I hereby certify that this New Application Transmittal and the documents referred to as enclosed therein are being deposited with the U.S. Postal Service on this date **May 24, 2000** in an envelope as "Express Mail Post Office to Addressee" Mailing Label Number **EL 541 885 115 US** addressed to: **Box Patent Application**, Assistant Commissioner For Patents, Washington, D.C. 20231.

Mary Ellen Waite
Mary Ellen Waite

1. Type Of Application

This new application is for a(n)

☒ Original (nonprovisional)

2. Papers Enclosed That Are Required For Filing Date Under 37 C.F.R. § 1.53(b) (Regular) or 37 C.F.R. § 1.153

(Design) Application

479 Pages of Specification

7 Pages of Claims

1 Page of Abstract

37 Sheets of Informal Drawings

3. Declaration

☒ Enclosed☒ Unexecuted.

4. Inventorship Statement

The inventorship for all the claims in this application is:

☒ the same

5. Language

☒ English

6. Assignment

☒ An unexecuted assignment of the invention to **Third Wave Technologies, Inc.** is attached.☒ Form PTO-1595 will follow.

7. Fee Calculation (37 C.F.R. § 1.16)

☒ Regular application

CLAIMS AS FILED

Number Filed	Number Extra	Rate	Basic Fee - \$760.00 (37 C.F.R. § 1.16(a))
Total Claims (37 C.F.R. § 1.16(c))	57 - 20 =	37 × \$18.00 =	\$666.00
Independent Claims (37 C.F.R. § 1.16(b))	2 - 3 =	0 × \$78.00 =	\$0.00
Multiple Dependent Claim(s), if any (37 C.F.R. § 1.16(d))	+ \$260.00 =		\$0.00

Filing Fee Calculation

\$1426.00

8. Small Entity Statement(s)

☒ An unexecuted Verified Statement(s) that this is a filing by a small entity under 37 C.F.R. §§ 1.9 and 1.27 is(are) attached.

Filing Fee Calculation (50% of above)

\$713.00

9. **Fee Payment Being Made At This Time**

☒ Enclosed

☒ basic filing fee

\$713.00

Total Fees Enclosed

\$713.00

10. **Method of Payment of Fees**

☒ Check in the amount of \$713.00

11. **Authorization To Charge Additional Fees and Credit Overpayment**

☒ The Commissioner is hereby authorized to charge payment of any fees associated with this communication or credit any overpayment to Deposit Account No.: **08-1290**. An originally executed duplicate of this transmittal is enclosed for this purpose.


12. **Power of Attorney by Assignee**

☒ Enclosed unexecuted

13. **Return Receipt Postcard**

☒ Enclosed

Dated: May 24, 2000


David A. Casimir
Registration No.: 42,395

MEDLEN & CARROLL, LLP
220 Montgomery Street, Suite 2200
San Francisco, California 94104
(415) 705-8410

☒ **Statement Where No Further Pages Added**

☒ This transmittal ends with this page.

Applicant / Patentee: Wu-Po Ma *et al.*
For: **Improved Enzymes For The Detection Of Specific Nucleic Acid Sequences**

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
(37 CFR § 1.9(c) - SMALL BUSINESS CONCERN)

I hereby declare that I am an official of the small business concern empowered to act on behalf of the concern identified below:

Third Wave Technologies, Inc.
502 South Rosa Road , Madison, WI 53719

I hereby declare that the above-identified small business concern qualifies as a small business concern as defined in 13 CFR §§ 121.3-18, and reproduced in 37 CFR § 1.9(d), for purposes of paying reduced fees under §§ 41(a) and (b) of Title 35, United States Code, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention entitled **Improved Enzymes For The Detection Of Specific Nucleic Acid Sequences** by inventors named **Wu-Po Ma *et al.***, described in the specification filed herewith.

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR § 1.28(b)).

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

Date: _____

By: _____

Name: _____

Title: _____

IMPROVED ENZYMES FOR THE DETECTION OF SPECIFIC NUCLEIC ACID SEQUENCES

This invention was made with United States Government support under Cooperative Agreement Number 70NANB7H3015, awarded by the National Institute of Standards and Technology (NIST). The United States Government has certain rights in the invention.

FIELD OF THE INVENTION

The present invention relates to novel enzymes designed for direct detection, characterization and quantitation of nucleic acids, particularly RNA. The present invention provides enzymes that recognize specific nucleic acid cleavage structures formed on a target RNA sequence and that cleave the nucleic acid cleavage structure in a site-specific manner to produce non-target cleavage products. The present invention provides enzymes having an improved ability to specifically cleave a DNA member of a complex comprising DNA and RNA nucleic acid strands.

BACKGROUND OF THE INVENTION

In molecular medicine, a simple and cost-effective method for direct and quantitative RNA detection would greatly facilitate the analysis of RNA viruses and the measurement of specific gene expression. Both of these issues are currently pressing problems in the field. Despite this need, few techniques have emerged that are truly direct. PCR-based detection assays require conversion of RNA to DNA by reverse transcriptase before amplification, introducing a variable that can compromise accurate quantification. Furthermore, PCR and other methods based on exponential amplification (*e.g.*, NASBA) require painstaking containment measures to avoid cross-contamination, and have difficulty distinguishing small differences (*e.g.*, 2 to 3-fold) in quantity. Other tests that directly examine RNA suffer from a variety of drawbacks, including time consuming autoradiography steps (*e.g.*, RNase protection assays), or overnight reaction times (*e.g.*, branched DNA assays). With over 1.5 million viral load measurements being performed in the U.S. every year, there is clearly an enormous potential for an inexpensive, rapid, high-throughput system for the quantitative

measurement of RNA.

Techniques for direct, quantitative detection of mRNA are vital for monitoring expression of a number of different genes. In particular, levels of cytokine expression (*e.g.*, interleukins and lymphokines) are being exploited as clinical measures of immune response in the progression of a wide variety of diseases (Van Deuren *et al.*, J. Int. Fed. Clin. Chem., 5:216 [1993], Van Deuren *et al.*, J. Inf. Dis., 169:157 [1994], Perenboom *et al.*, Eur. J. Clin. Invest., 26:159 [1996], Guidotti *et al.*, Immunity 4:25 [1996]) as well as in monitoring transplant recipients (Grant *et al.*, Transplantation 62:910 [1996]). Additionally, the monitoring of viral load and identification of viral genotype have great clinical significance for individuals suffering viral infections by such pathogens as HIV or Hepatitis C virus (HCV). There is a high correlation between viral load (*i.e.*, the absolute number of viral particles in the bloodstream) and time to progression to AIDS (Mellors *et al.*, Science 272:1167 [1996], Saag *et al.*, Nature Medicine 2:625 [1996]). For that reason, viral load, as measured by quantitative nucleic acid based testing, is becoming a standard monitoring procedure for evaluating the efficacy of treatment and the clinical status of HIV positive patients. It is thought to be essential to reduce viral load as early in the course of infection as possible and to evaluate viral levels on a regular basis. In the case of HCV, viral genotype has great clinical significance, with correlations to severity of liver disease and responsiveness to interferon therapy. Furthermore, because HCV cannot be grown in culture, it is only by establishing correlations between characteristics like viral genotype and clinical outcome that new antiviral treatments can be evaluated.

While the above mentioned methods have been serviceable for low throughput, research applications, or for limited clinical application, it is clear that large scale quantitative analysis of RNA readily adaptable to any genetic system will require a more innovative approach. An ideal direct detection method would combine the advantages of the direct detection assays (*e.g.*, easy quantification and minimal risk of carry-over contamination) with the specificity provided by a dual oligonucleotide hybridization assay.

Many of the methods described above rely on hybridization alone to distinguish a target molecule from other nucleic acids. Although some of these methods can be highly

sensitive, they often cannot quantitate and distinguish closely related mRNAs accurately, especially such RNAs expressed at different levels in the same sample. While the above-mentioned methods are serviceable for some purposes, a need exists for a technology that is particularly adept at distinguishing particular RNAs from closely related molecules.

5

SUMMARY OF THE INVENTION

The present invention relates to novel enzymes designed for direct detection, characterization and quantitation of nucleic acids, particularly RNA. The present invention provides enzymes that recognize specific nucleic acid cleavage structures formed on a target RNA sequence and that cleave the nucleic acid cleavage structure in a site-specific manner to produce non-target cleavage products. The present invention provides enzymes having an improved ability to specifically cleave a DNA member of a complex comprising DNA and RNA nucleic acid strands.

10

For example, the present invention provides DNA polymerases that are altered in structure relative to the native DNA polymerases, such that they exhibit altered (*e.g.*, improved) performance in detection assays based on the cleavage of a structure comprising nucleic acid (*e.g.*, RNA). In particular, the altered polymerases of the present invention exhibit improved performance in detection assays based on the cleavage of a DNA member of a cleavage structure (*e.g.*, an invasive cleavage structure) that comprises an RNA target strand.

15

The improved performance in a detection assay may arise from any one of, or a combination of several improved features. For example, in one embodiment, the enzyme of the present invention may have an improved rate of cleavage (k_{cat}) on a specific targeted structure, such that a larger amount of a cleavage product may be produced in a given time span. In another embodiment, the enzyme of the present invention may have a reduced activity or rate in the cleavage of inappropriate or non-specific structures. For example, in certain embodiments of the present invention, one aspect of improvement is that the differential between the detectable amount of cleavage of a specific structure and the detectable amount of cleavage of any alternative structures is increased. As such, it is within the scope of the present invention to provide an enzyme having a reduced rate of cleavage of a specific target structure compared to the rate of the native enzyme, and having a further

20

25

30

reduced rate of cleavage of any alternative structures, such that the differential between the detectable amount of cleavage of the specific structure and the detectable amount of cleavage of any alternative structures is increased. However, the present invention is not limited to enzymes that have an improved differential.

5 In a preferred embodiment, the enzyme of the present invention is a DNA polymerase having an altered nuclease activity as described above, and also having altered synthetic activity, compared to that of any native DNA polymerase from which the enzyme has been derived. It is especially preferred that the DNA polymerase is altered such that it exhibits reduced synthetic activity as well as improved nuclease activity on RNA targets, compared to
10 that of the native DNA polymerase. Enzymes and genes encoding enzymes having reduced synthetic activity have been described (*See e.g.*, Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999], Lyamichev *et al.*, Prot. Natl. Acad. Sci., 96:6143 [1999], US. Patents 5,541,311, 5,614,402, 5,795,763 and U.S. Patent application Ser. No. 08/758,314, incorporated herein by reference in their entireties). The present invention contemplates combined modifications, such that the resulting 5' nucleases are without interfering synthetic activity, and have improved performance in RNA detection assays.

The present invention contemplates a DNA sequence encoding a DNA polymerase altered in sequence relative to the native sequence, such that it exhibits altered nuclease activity from that of the native DNA polymerase. For example, in one embodiment, the DNA
20 sequence encodes an enzyme having an improved rate of cleavage (k_{cat}) on a specific targeted structure, such that a larger amount of a cleavage product may be produced in a given time span. In another embodiment, the DNA encodes an enzyme having a reduced activity or rate in the cleavage of inappropriate or non-specific structures. In certain embodiments, one aspect of improvement is that the differential between the detectable amount of cleavage of a
25 specific structure and the detectable amount of cleavage of any alternative structures is increased. It is within the scope of the present invention to provide a DNA encoding an enzyme having a reduced rate of cleavage of a specific target structure compared to the rate of the native enzyme, and having a further reduced rate of cleavage of any alternative structures, such that the differential between the detectable amount of cleavage of the specific
30 structure and the detectable amount of cleavage of any alternative structures is increased.

However, the present invention is not limited to polymerases that have an improved differential.

In a preferred embodiment, the DNA sequence encodes a DNA polymerase having the altered nuclease activity described above, and also having altered synthetic activity, compared to that of any native DNA polymerase from which the improved enzyme is derived. It is especially preferred that the encoded DNA polymerase is altered such that it exhibits reduced synthetic activity as well as improved nuclease activity on RNA targets, compared to that of the native DNA polymerase.

It is not intended that the invention be limited by the nature of the alteration required to introduce altered nuclease activity. Nor is it intended that the invention be limited by the extent of either the alteration, or in the improvement observed. If the polymerase is also altered so as to be synthesis modified, it is not intended that the invention be limited by the polymerase activity of the modified or unmodified protein, or by the nature of the alteration to render the polymerase synthesis modified.

The present invention contemplates structure-specific nucleases from a variety of sources, including, but not limited to, mesophilic, psychrophilic, thermophilic, and hyperthermophilic organisms. The preferred structure-specific nucleases are thermostable. Thermostable structure-specific nucleases are contemplated as particularly useful in that they allow the INVADER assay (*See e.g.*, U.S. Pat. Nos. 5,846,717, 5,985,557, 5,994,069, and 6,001,567 and PCT Publications WO 97/27214 and WO 98/42873, incorporated herein by reference in their entireties) to be operated near the melting temperature (T_m) of the downstream probe oligonucleotide, so that cleaved and uncleaved probes may cycle on and off the target during the course of the reaction. In one embodiment, the thermostable structure-specific enzymes are thermostable 5' nucleases that are selected from the group comprising altered polymerases derived from the native polymerases of *Thermus* species, including, but not limited to, *Thermus aquaticus*, *Thermus flavus*, *Thermus thermophilus*, *Thermus filiformus*, and *Thermus scotoductus*. However, the invention is not limited to the use of thermostable 5' nucleases. For example, certain embodiments of the present invention utilize short oligonucleotide probes that may cycle on and off of the target at low temperatures, allowing the use of non-thermostable enzymes.

In some preferred embodiments, the present invention provides a composition comprising an enzyme, wherein the enzyme comprises a heterologous functional domain, wherein the heterologous functional domain provides altered (*e.g.*, improved) functionality in a nucleic acid cleavage assay. The present invention is not limited by the nature of the nucleic acid cleavage assay. For example, nucleic acid cleavage assays include any assay in which a nucleic acid is cleaved, directly or indirectly, in the presence of the enzyme. In certain preferred embodiments, the nucleic acid cleavage assay is an invasive cleavage assay. In particularly preferred embodiments, the cleavage assay utilizes a cleavage structure having at least one RNA component. In another particularly preferred embodiment, the cleavage assay utilizes a cleavage structure having at least one RNA component, wherein a DNA member of the cleavage structure is cleaved.

In some preferred embodiments, the enzyme comprises a 5' nuclease or a polymerase. In certain preferred embodiments, the 5' nuclease comprises a thermostable 5' nuclease. In other preferred embodiments, the polymerase is altered in sequence relative to a naturally occurring sequence of a polymerase such that it exhibits reduced DNA synthetic activity from that of the naturally occurring polymerase. In certain preferred embodiments, the polymerase comprises a thermostable polymerase (*e.g.*, a polymerase from a *Thermus* species including, but not limited to, *Thermus aquaticus*, *Thermus flavus*, *Thermus thermophilus*, *Thermus filiformus*, and *Thermus scotoductus*).

The present invention is not limited by the nature of the altered functionality provided by the heterologous functional domain. Illustrative examples of alterations include, but are not limited to, enzymes where the heterologous functional domain comprises an amino acid sequence (*e.g.*, one or more amino acids) that provides an improved nuclease activity, an improved substrate binding activity and/or improved background specificity in a nucleic acid cleavage assay.

The present invention is not limited by the nature of the heterologous functional domain. For example, in some embodiments, the heterologous functional domain comprises two or more amino acids from a polymerase domain of a polymerase (*e.g.*, introduced into the enzyme by insertion of a chimeric functional domain or created by mutation). In certain

preferred embodiment, at least one of the two or more amino acids is from a palm or thumb region of the polymerase domain. The present invention is not limited by the identity of the polymerase from which the two or more amino acids are selected. In certain preferred embodiments, the polymerase comprises *Thermus thermophilus* polymerase. In particularly preferred embodiments, the two or more amino acids are from amino acids 300-650 of SEQ ID NO:267.

The novel enzymes of the invention may be employed for the detection of target DNAs and RNAs including, but not limited to, target DNAs and RNAs comprising wild type and mutant alleles of genes, including, but not limited to, genes from humans, other animal, or plants that are or may be associated with disease or other conditions. In addition, the enzymes of the invention may be used for the detection of and/or identification of strains of microorganisms, including bacteria, fungi, protozoa, ciliates and viruses (and in particular for the detection and identification of viruses having RNA genomes, such as the Hepatitis C and Human Immunodeficiency viruses). For example, the present invention provides methods for cleaving a nucleic acid comprising providing: an enzyme of the present invention and a substrate nucleic acid; and exposing the substrate nucleic acid to the enzyme (*e.g.*, to produce a cleavage product that may be detected).

In one embodiment, the present invention provides a thermostable 5' nuclease having an amino acid sequence selected from the group comprising SEQ ID NOS:75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 106, 109, 112, 115, 118, 121, 124, 127, 130, 133, 136, 139, 142, 145, 148, 150, 153, 157, 160, 163, 166, 169, 172, 175, 178, 181, 184, 187, 190, 200, 202, 204, 206, 212, 214, 216, 218, 221, 226, 228, 230, 232, 234, 236, 239, 241, 243, 251, 259, 261, 263, and 265. In another embodiment, the 5' nuclease is encoded by a DNA sequence selected from the group comprising of SEQ ID NOS:74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 105, 108, 111, 114, 117, 120, 123, 126, 129, 132, 135, 138, 141, 144, 147, 149, 152, 156, 159, 162, 165, 168, 171, 174, 177, 180, 183, 186, 189, 199, 201, 203, 205, 211, 213, 215, 217, 220, 225, 227, 229, 231, 233, 235, 238, 240, 242, 250, 258, 260, 262, and 264.

The present invention also provides a recombinant DNA vector comprising DNA having a nucleotide sequence encoding a 5' nuclease, the nucleotide sequence selected from

the group comprising SEQ ID NOS:74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 105, 108, 111, 114, 117, 120, 123, 126, 129, 132, 135, 138, 141, 144, 147, 149, 152, 156, 159, 162, 165, 168, 171, 174, 177, 180, 183, 186, 189, 199, 201, 203, 205, 211, 213, 215, 217, 220, 225, 227, 229, 231, 233, 235, 238, 240, 242, 250, 258, 260, 262, and 264. In a preferred embodiment, the invention provides a host cell transformed with a recombinant DNA vector comprising DNA having a nucleotide sequence encoding a structure-specific nuclease, the nucleotide selected from the group comprising sequence SEQ ID NOS:74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 105, 108, 111, 114, 117, 120, 123, 126, 129, 132, 135, 138, 141, 144, 147, 149, 152, 156, 159, 162, 165, 168, 171, 174, 177, 180, 183, 186, 189, 199, 201, 203, 205, 211, 213, 215, 217, 220, 225, 227, 229, 231, 233, 235, 238, 240, 242, 250, 258, 260, 262, and 264. The invention is not limited by the nature of the host cell employed. The art is well aware of expression vectors suitable for the expression of nucleotide sequences encoding structure-specific nucleases that can be expressed in a variety of prokaryotic and eukaryotic host cells. In a preferred embodiment, the host cell is an *Escherichia coli* cell.

The present invention provides a method of altering 5' nuclease enzymes relative to native 5' nuclease enzymes, such that they exhibit improved performance in detection assays based on the cleavage of a structure comprising RNA. In particular, the altered 5' nucleases produced by the method of the present invention exhibit improved performance in detection assays based on the cleavage of a DNA member of a cleavage structure (*e.g.*, an invasive cleavage structure) that comprises an RNA target strand. The improved 5' nucleases resulting from the methods of the present invention may be improved in any of the ways discussed herein. Examples of processes for assessing improvement in any candidate enzyme are provided.

For example, the present invention provides methods for producing an altered enzyme with improved functionality in a nucleic acid cleavage assay comprising: providing an enzyme and a nucleic acid test substrate; introducing a heterologous functional domain into the enzyme to produce an altered enzyme; contacting the altered enzyme with the nucleic acid test substrate to produce cleavage products; and detecting the cleavage products. In some embodiments, the introduction of the heterologous functional domain comprises mutating one

or more amino acids of the enzyme. In other embodiments, the introduction of the heterologous functional domain into the enzyme comprises adding a functional domain from a protein (e.g., another enzyme) into the enzyme (e.g., substituting functional domains by removing a portion of the enzyme sequence prior to adding the functional domain of the protein). In preferred embodiments, the nucleic acid test substrate comprises a cleavage structure. In particularly preferred embodiment, the cleavage structure comprises an RNA target nucleic acid. In yet other preferred embodiments, the cleavage structure comprises an invasive cleavage structure.

The present invention also provides nucleic acid treatment kits. One preferred embodiment is a kit comprising a composition comprising at least one improved 5' nuclease. Another preferred embodiment provides a kit comprising: a) a composition comprising at least one improved 5' nuclease; and b) an INVADER oligonucleotide and a signal probe oligonucleotide. In some embodiments of the kits of the present invention, the improved 5' nuclease is derived from a DNA polymerase from a eubacterial species. In further embodiments, the eubacterial species is a thermophile. In still further embodiments, the thermophile is of the genus *Thermus*. In still further embodiments, the thermophile is selected from the group consisting of *Thermus aquaticus*, *Thermus flavus*, *Thermus thermophilus*, *Thermus filiformus*, and *Thermus scotoductus*. In preferred embodiments, the improved 5' nuclease is encoded by DNA selected from the group comprising SEQ ID NOS:74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 105, 108, 111, 114, 117, 120, 123, 126, 129, 132, 135, 138, 141, 144, 147, 149, 152, 156, 159, 162, 165, 168, 171, 174, 177, 180, 183, 186, 189, 199, 201, 203, 205, 211, 213, 215, 217, 220, 225, 227, 229, 231, 233, 235, 238, 240, 242, 250, 258, 260, 262, and 264. In yet other preferred embodiments, the kits further comprise reagents for detecting a nucleic acid cleavage product. In further preferred embodiments, the reagents for detecting a cleavage product comprise oligonucleotides for use in a subsequent invasive cleavage reaction (See e.g., U.S. Patent No. 5,994,069). In particularly preferred embodiments, the reagents for the subsequent invasive cleavage reaction comprise a probe labeled with moieties that produce a fluorescence resonance energy transfer (FRET) effect.

DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic representation of sequential invasive cleavage reactions. In step A, an upstream INVADER oligonucleotide and a downstream probe combine with a target nucleic acid strand to form a cleavage structure. In step B, the portion of the cleaved signal probe from A combines with a second target nucleic acid strand and a labeled signal probe to form a second cleavage structure. In step C, cleavage of the labeled second cleavage structure yields a detectable signal.

Figure 2 shows schematic representations of several examples of invasive cleavage structures comprising RNA target strands (SEQ ID NO:9). Panel A depicts an INVADER oligonucleotide (SEQ ID NO:10) and probe (SEQ ID NO:11). Panel B depicts an INVADER oligonucleotide (SEQ ID NO:12) and probe (SEQ ID NO:11). Panel C depicts an INVADER oligonucleotide (SEQ ID NO:13) and probe (SEQ ID NO:13). Panel D depicts an INVADER oligonucleotide (SEQ ID NO:13) and probe (SEQ ID NO:14).

Figure 3 shows schematic representations of two examples of structures that are not invasive cleavage structures labelled SEQ ID NOs:269-274.

Figure 4 shows a schematic representation of a configuration of invasive cleavage that is useful for detection of target sequence variations. In A, an invasive cleavage structure having overlap between the two probes is formed, and the arrow indicates that it is cleavable by the enzymes of the present invention. In B, variation of the target sequence removes a region of complementarity to the downstream probe and eliminates the overlap. The absence of an arrow in panel B indicates a reduced rate of cleavage of this structure compared to that diagrammed in panel A.

Figure 5 shows a diagram of the X-ray structure of a ternary complex of KlenTaq1 with primer/template DNA in the polymerizing mode determined by Li *et al.* (Li *et al.*, Protein Sci., 7:1116 [1998]). Without intending to represent precise borders between features of the physical form, the portions referred to in the text as the "fingers", "thumb" and "palm" regions are loosely indicated by the circle, rectangle, and oval, respectively.

Figure 6 shows a schematic diagram of the DNA polymerase gene from *Thermus aquaticus*. Restriction sites used in these studies are indicated above. The approximate regions encoding various structural or functional domains of the protein are indicated by

double-headed arrows, below.

Figure 7 shows a schematic diagram of the chimeric constructs comprising portions of the TaqPol gene and the TthPol gene. Open and shaded boxes denote TaqPol and TthPol sequences, respectively. The numbers correspond to the amino acid sequence of TaqPol. The 5' nuclease and polymerase domains of TaqPol and the palm, thumb, and fingers regions of the polymerase domain are indicated. The abbreviations for the restrictions sites used for recombination are as follows: E, EcoRI; N, NotI; Bs, BstBI; D, NdeI; B, BamHI; and S, SalI.

Figure 8 is a comparison of the nucleotide structure of the polymerase genes isolated from *Thermus aquaticus* (SEQ ID NO:1), *Thermus flavus* (SEQ ID NO:2) and *Thermus thermophilus* (SEQ ID NO:266); the consensus sequence (SEQ ID NO:7) is shown at the top of each row.

Figure 9 is a comparison of the amino acid sequence of the polymerase isolated from *Thermus aquaticus* (SEQ ID NO:4), *Thermus flavus* (SEQ ID NO:5), and *Thermus thermophilus* (SEQ ID NO:267); the consensus sequence (SEQ ID NO:8) is shown at the top of each row.

Figure 10 shows the sequences and proposed structures of substrates for the invasive signal amplification reaction with human IL-6 RNA target strand (SEQ ID NO:17) and upstream probe (SEQ ID NO:15). The cleavage site of the downstream probe (SEQ ID NO:16) is indicated by an arrow. Sequence of the IL-6 DNA target strand (SEQ ID NO:18) is shown below.

Figure 11 shows the image generated by a fluorescence imager showing the products of invasive cleavage assays using the indicated enzymes, and the IL-6 substrate of Figure 10 having either a DNA target strand (A) or an RNA target strand (B).

Figure 12 compares the cycling cleavage activities of Taq DN RX HT, Tth DN RX HT, and Taq-Tth chimerical enzymes with IL-6 substrate having an RNA target strand.

Figure 13 shows a comparison of the amino acid sequences of the BstI-BamHI fragments of TaqPol (SEQ ID NO:19) and TthPol (SEQ ID NO:20). Pairs of similar amino acids are shaded with light gray. Aligned amino acids that have a charge difference are shaded with dark gray. The numbers correspond to the amino acid sequence of TaqPol.

Amino acids of TaqPol changed to the corresponding amino acids of TthPol by site-directed mutagenesis are indicated by (+).

Figure 14 compares the cycling cleavage activities of Taq DN RX HT, Taq-Tth chimerical enzymes, and chimerical enzymes having the indicated additional amino acid modifications, with IL-6 substrate having an RNA target strand.

Figure 15 compares the cycling cleavage activities of Taq DN RX HT, Tth DN RX HT, and Taq DN RX HT having the indicated amino acid modifications, with IL-6 substrate having an RNA target strand.

Figure 16 compares polymerization activities of TaqPol, TthPol, and Taq-Tth chimerical enzymes, and TaqPol having the indicated amino acid modifications.

Figure 17 shows a diagram of the X-ray structure of a ternary complex of KlenTaqI with primer/template DNA in the polymerizing mode determined by Li *et al.* (Li *et al.*, Protein Sci., 7:1116 [1998]). Amino acids G418 and E507 are indicated.

Figures 18 A-D show schematic diagrams of examples of substrates that may be used to measure various cleavage activities of enzymes. The substrates may be labeled, for example, with a fluorescent dye and a quenching moiety for FRET detection, as shown, to facilitate detection and measurement. The substrates of 18A and 18B are invasive cleavage structures having RNA and DNA target strands, respectively. 18C shows an example of an X-structure, and 18D shows an example of a hairpin structure, both of which may be used to assess the activity of enzymes on alternative structures that may be present in invasive cleavage reactions.

Figure 19 shows schematic diagrams of chimeric constructs comprising portions of the TaqPol gene and the TthPol gene. Open and shaded boxes denote TaqPol and TthPol sequences, respectively. The chimeras also include the DN, RX, and HT modifications. A table compares the cleavage activity of each protein on the indicated cleavage substrates.

Figure 20A shows a schematic diagram for an RNA containing invasive cleavage substrate. The 5' end of the target molecule (SEQ ID NO:27) is modified with biotin and blocked with streptavidin as described. The downstream probe (SEQ ID NO:26) with cleavage site is also shown. Panels B-D show analysis of the properties of the Taq DN RX HT G418K/E507Q mutant in cleavage of the shown substrate under conditions of varying

reaction temperature, KCl concentration, and MgSO₄ concentration.

Figure 21 shows schematic diagrams for model substrates used to test enzymes for invasive cleavage activity. The molecule shown in 21A provides a DNA target strand (SEQ ID NO:28), while the model shown in 21B provides an RNA containing target strand (SEQ ID NO:27). Both 21A and B show downstream probe SEQ ID NO:26.

Figure 22 shows schematic diagrams for model substrates used to test enzymes for cleavage activity on alternative, non-invasive structures.

Figure 23 shows a schematic diagram for a model substrate used to test enzymes for invasive cleavage activity.

Figure 24 shows schematic diagrams for a model substrate used to test enzymes for invasive cleavage activity on an RNA or DNA target strands.

Figure 25 compares the cycling cleavage activities of Tth DN RX HT, Taq 2M, TfiPol, Tsc Pol, and Tfi and Tsc-derived mutant enzymes.

GENERAL DESCRIPTION OF THE INVENTION

The INVADER technology (*See e.g.*, U.S. Pat. Nos. 5,846,717, 5,985,557, 5,994,069, and 6,001,567 and PCT Publications WO 97/27214 and WO 98/42873, incorporated herein by reference in their entireties) provides a signal amplification system that can be applied to the detection and quantitation of specific nucleic acid sequences, including single point mutations and similar variants of mRNA. Further, because this technology does not rely exclusively on allele-specific hybridization, it is well suited for quantitating closely related RNAs in the same sample. The present invention provides improved enzymes and methods for creating enzymes for the INVADER assay-based detection of nucleic acids, particularly RNA nucleic acids. The present invention also provides kits for the performance of INVADER assays using the improved enzymes of the present invention.

The INVADER technology was developed for quantitative detection of DNA and RNA, without prior amplification of the target nucleic acid (Lyamichev *et al.*, Nat. Biotechnol., 17:292 [1999]). In addition to its use for the quantitative measurement of specific nucleic acid sequences, high specificity provides the capability of detecting single base changes. The basis of the INVADER technology is the cleavage of DNA and RNA

molecules at specific locations in response to structure rather than sequence. Cleavage is typically catalyzed by a 5' nuclease enzyme. The 5' nuclease enzymes recognize a precise structure that is formed when two oligonucleotide probes, an upstream INVADER probe and a downstream signal probe, hybridize in tandem to a nucleic acid target to generate the substrate complex (Figure 1A). The high specificity of the INVADER technology arises from combining sequence-specific probe hybridization with structure-specific enzymatic cleavage. The substrate complex contains a feature that is important for precise enzyme recognition: an overlap between the hybridized oligonucleotides. To form an invasive structure, the 3' end of the upstream INVADER oligonucleotide must overlap with the hybridized region of the signal probe by at least one base (Lyamichev *et al.*, Nat. Biotechnol., 17:292 [1999]). This overlap may be created by a duplication of sequence between the 3' portion of the upstream INVADER oligonucleotide and the 5' portion of the target-complementary region of the downstream probe oligonucleotide. The region of sequence so duplicated may be as small as a single base. Regardless of the length of the duplicated sequence (*i.e.*, the overlap) the 3' terminal base of the upstream INVADER oligonucleotide need not be complementary to the target strand, and may be any nucleotide. In some embodiments, this terminal nucleotide may be replaced by a moiety having chemical features similar to a nucleotide such as a nucleotide analog or an organic ring compound (*See e.g.*, U.S. Pat. No. 5,985,557). In an alternative embodiment, the overlap need not involve any duplication of sequence between the target-complementary regions of the two probes (Lyamichev *et al.*, Nat. Biotechnol., 17:292 [1999] and U.S. Patent 5,985,557). In this embodiment, the INVADER and signal probes have regions complementary to adjacent regions of the target that are contiguous and that do not overlap. When no sequence is shared, the 3' end of the upstream INVADER oligonucleotide includes at least one additional nucleotide or nucleotide-like analog that is not complementary to the target strand (Lyamichev *et al.*, Nat. Biotechnol., 17:292 [1999]). This can be referred to as a physical overlap, in contrast to a sequence overlap. An overlap of either type will satisfy the requirement for overlap that is the hallmark of the invasive cleavage of the INVADER assay. Several of these embodiments are shown schematically in Figure 2. In contrast to the overlap configurations described above, if the probes have regions

complementary to adjacent regions of the target that are contiguous and that do not overlap, and the 3' end of the upstream oligonucleotide does not have any additional base or moiety, the invasive structure is not formed (Figure 3A). Even the presence of one or more additional bases on the 5' end of the downstream oligonucleotide that are not complementary to the target strand will not create the requisite overlap. This latter structure (Figure 3B), as is described in U.S. 5,874,283 is not an "invasive cleavage," although such structures find use in certain embodiments of the present invention.

Some 5' nucleases may not require an upstream oligonucleotide to be active in a cleavage reaction. Although cleavage may be slower without the upstream oligonucleotide, it may still occur (Lyamichev *et al.*, Science 260:778 [1993], Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999]). When a DNA strand is the template or target strand to which probe oligonucleotides are hybridized, the 5' nucleases derived from DNA polymerases and some flap endonucleases (FENs), such as that from *Methanococcus jannaschii*, can cleave quite well without an upstream oligonucleotide providing an overlap (Lyamichev *et al.*, Science 260:778 [1993], Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999], and US Patent No. 5,843,669, herein incorporated by reference in its entirety). Other FENs, such as those from *Archeoglobus fulgidus* (Afu) and *Pyrococcus furiosus* (Pfu), cleave an overlapped structure on a DNA target at so much greater a rate than they do a non-overlapping structure (*i.e.*, either missing the upstream oligonucleotide or having a non-overlapping upstream oligonucleotide) that they can be viewed as having an essentially absolute requirement for the overlap (Lyamichev *et al.*, Nat. Biotechnol., 17:292 [1999], Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999]). When an RNA target is hybridized to DNA oligonucleotide probes to form a cleavage structure, many FENs cleave the downstream DNA probe poorly, regardless of the presence of an overlap. On such an RNA-containing structure, the 5' nucleases derived from DNA polymerases have a strong requirement for the overlap, and are essentially inactive in its absence.

Performing the INVADER assay under conditions that have a tight requirement for an overlap (*e.g.*, using the *Afu* FEN for DNA target detection or the 5' nuclease of *Tth* DNA polymerase for RNA target detection) provides a superior means of detecting single nucleotide or other sequence variations. In one embodiment, the signal probe is selected such that the

target base suspected of varying is positioned at the 5' end of the target-complementary region of this probe. The upstream INVADER oligonucleotide is positioned to provide a single base of overlap. If the target and the signal probe are complementary at the base in question, the overlap forms and cleavage can occur. This embodiment is diagrammed in Figure 4A.

5 However, if the target does not complement the probe at this position, that base in the probe becomes part of a non-complementary 5' arm, no overlap between the probes exists, and cleavage is suppressed. This embodiment is diagrammed in Figure 4B. In any of the aforementioned embodiments, the downstream probe may optionally include a region that is not complementary to the target. In a preferred embodiment, this non target-complementary region is on the 5' end of the probe and produces an unpaired 5' flap when the signal probe is hybridized to the target. Upon cleavage by a CLEAVASE enzyme, a released 5' flap can be incorporated into a subsequent INVADER reaction for further amplification of the signal (See e.g., U.S. Patent 5,994,069 and PCT Publication WO 98/42873, incorporated herein by reference in their entireties). One way it may be used is as an INVADER oligonucleotide, which may combine with a provided secondary target and a secondary probe. Upon hybridization of the 5' flap released by the CLEAVASE enzyme in the first invasive cleavage reaction, a secondary invasive structure complex is completed, so that it may be recognized by the CLEAVASE enzyme and the secondary probe oligonucleotide may be cleaved (Kwiatkowski *et al.*, Molec. Diagn., 4:353 [1999]), Figure 1.

10
15
20
25
30
INVADER assays often use thermostable CLEAVASE enzymes, allowing reactions to be operated near the melting temperature (T_m) of the downstream probe oligonucleotide, so that cleaved and uncleaved probes cycle on and off the target during the course of the reaction. In a preferred embodiment, a longer INVADER oligonucleotide may not readily cycle. Each time a full-length probe binds to the target in the presence of the INVADER oligonucleotide it can be cleaved, resulting in an accumulation of cleavage product that is both highly specific for the sequence being detected, and that is generally proportional with respect to both time and target concentration. The target is generally the limiting component in an invasive cleavage, since the INVADER and signal probe oligonucleotides are generally supplied in molar excess. In a second linked invasive cleavage, it is the component created in the first cleavage reaction (e.g., a released 5' flap) that is limiting. When two such cleavage

reactions are performed sequentially, the signal from the composite reaction accumulates linearly with respect to the amount of target nucleic acid while the reaction sequence results in a tremendous increase in signal amplification (Kwiatkowski *et al.*, Molec. Diagn., 4:353 [1999]).

Several of the 5' nuclease domains of eubacterial Pol A DNA polymerases and structurally homologous DNA repair proteins, called flap endonucleases (FENs) can function to cleave the secondary structure formed between the INVADER and signal probe oligonucleotides (Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999], Xu *et al.*, J. Biol. Chem., published online as 10.1074/jbc.M909135199 at www.jbc.org/pips/pips.2.shtml, May 9, 2000). Both classes of enzymes contain a putative helix-hairpin-helix (HhH) DNA binding motif important for sequence independent, structure-based recognition of DNA (Doherty *et al.*, Nucl. Acid. Res., 24:2488 [1996]). This type of DNA binding motif is suitable for assays performed on DNA targets, but can be problematic for assays with RNA targets, resulting in lower assay sensitivity. New enzymes having improved recognition of the invasive cleavage structure formed on an RNA target strand would vastly improve the performance of the INVADER assay in the detection and quantitation of RNA targets.

A number of enzyme improvements related to 5' nucleases and DNA polymerases have been described. For example, DNA polymerases having altered 5' nuclease activity, or lacking 5' nuclease activity altogether have been described (U.S. Patents 5,466,591 and 5,795,762, each of which is incorporated herein by reference in its entirety). These patents relate to thermostable DNA polymerases that exhibit a different level of 5' to 3' exonuclease activity than their respective native polymerases. In some embodiments, particular conserved amino acid domains in thermostable DNA polymerases are mutated or deleted to alter the 5' to 3' exonuclease activity of the polymerases.

DNA polymerases altered relative to the native polymerases such that they exhibit altered DNA synthetic activity have been described (Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999], Lyamichev *et al.*, Proc. Natl. Acad. Sci., 96:6143 [1999], US. Patents 5,541,311, 5,614,402, 5,795,763 and U.S. Patent application Ser. No. 08/758,314, incorporated herein by reference in their entireties). In preferred embodiments, these DNA polymerases are altered

such that they exhibit reduced synthetic activity compared to that of the native DNA polymerase. In this respect, enzymes have been created that are predominantly 5' nucleases and are capable of cleaving nucleic acids in a structure-specific manner in the absence of interfering synthetic activity. The alterations made in these polymerases were not selected with respect to their effect of the cleavage of structures comprising RNA.

DNA polymerases having the ability to use RNA as a template strand, known as reverse transcriptases, are usually associated with an RNase activity that specifically cleaves RNA basepaired in a heteroduplex with a DNA strand. Such RNase activity is generally termed RNase H. Altered reverse transcriptases that have this RNase H activity removed have been described (*See e.g.*, U.S. Patent 5,244,797, incorporated herein by reference in its entirety). This patent relates to a gene that encodes reverse transcriptase having DNA polymerase activity and little or no RNase H activity. The invention also relates to a method of producing cDNA from mRNA using the reverse transcriptase. This patent does not describe enzymes having improved ability to cleave a DNA member of a structure comprising DNA and RNA strands, nor does it relate to enzymes having improved performance in detection assays based on the cleavage of a DNA member of a structure that comprises an RNA target strand.

Thermostable RNase H enzymes have been described (*e.g.*, U.S. Patent Nos. 5,268,289, 5,459,055 and 5,500,370, incorporated herein by reference in their entireties). These thermostable enzymes cleave the RNA member of a heteroduplex comprising DNA and RNA strands. These patents do not describe enzymes having improved ability to cleave a DNA member of a structure comprising DNA and RNA strands, nor do they relate to enzymes having improved performance in detection assays based on the cleavage of a DNA member of an invasive structure that comprises an RNA target strand.

There remains a need for enzymes having an improved ability to cleave DNA members of structures comprising RNA and DNA strands. In particular, there remains a need for thermostable enzymes having improved performance in detection assays based on the cleavage of DNA members of invasive complexes comprising an RNA target strand.

DEFINITIONS

To facilitate an understanding of the present invention, a number of terms and phrases are defined below:

As used herein, the term "functional domain" refers to a region, or a part of a region, of a protein (*e.g.*, an enzyme) that provides one or more functional characteristic of the protein. For example, a functional domain of an enzyme may provide, directly or indirectly, one or more activities of the enzyme including, but not limited to, substrate binding capability and catalytic activity. A functional domain may be characterized through mutation of one or more amino acids within the functional domain, wherein mutation of the amino acid(s) alters the associated functionality (as measured empirically in an assay) thereby indicating the presence of a functional domain.

As used herein, the term "heterologous functional domain" refers to a protein functional domain that is not in its natural environment. For example, a heterologous functional domain includes a functional domain from one enzyme introduced into another enzyme. A heterologous functional domain also includes a functional domain native to an protein that has been altered in some way (*e.g.*, mutated, added in multiple copies, etc.). A heterologous functional domain may comprise a plurality of contiguous amino acids or may include two or more distal amino acids or amino acids fragments (*e.g.*, two or more amino acids or fragments with intervening, non-heterologous, sequence). Heterologous functional domains are distinguished from endogenous functional domains in that the heterologous amino acid(s) are joined to amino acid sequences that are not found naturally associated with the amino acid sequence in nature or are associated with a portion of a protein not found in nature.

As used herein, the term "altered functionality in a nucleic acid cleavage assay" refers to a characteristic of an enzyme that has been altered in some manner to differ from its natural state (*e.g.*, to differ from how it is found in nature). Alterations include, but are not limited to, addition of a heterologous functional domain (*e.g.*, through mutation or through creation of chimeric proteins). In some embodiments, the altered characteristic of the enzyme may be one that improves the performance of an enzyme in a nucleic acid cleavage assay.

Types of improvement include, but are not limited to, improved nuclease activity (*e.g.*, improved rate of reaction), improved substrate binding (*e.g.*, increased or decreased binding of certain nucleic acid species [*e.g.*, RNA or DNA] that produces a desired outcome [*e.g.*, greater specificity, improved substrate turnover, etc.]), and improved background specificity (*e.g.*, less undesired product is produced). The present invention is not limited by the nucleic cleavage assay used to test improved functionality. However, in some preferred embodiments of the present invention, an invasive cleavage assay is used as the nucleic acid cleavage assay. In certain particularly preferred embodiments, an invasive cleavage assay utilizing an RNA target is used as the nucleic acid cleavage assay.

The terms "signal oligonucleotide" and "signal probe" are used interchangeably, and refer to a member of a cleavage structure that is cleaved by a cleavage agent. In an "invasive cleavage structure," the signal probe may be referred to as the "downstream probe" in a complex in that, with the exception of its overlapping portion, it is largely 3' of an INVADER oligonucleotide within the complex. On the target strand, the hybridization site of a signal probe is said to be upstream of the hybridization site of an INVADER oligonucleotide.

The terms "INVADER" oligonucleotide or probe and "invasive" oligonucleotide or probe may be used interchangeably, and refer to a member of an invasive cleavage structure that hybridizes to a target nucleic acid strand and that overlaps the 5' end of a signal probe. In an invasive cleavage structure, the INVADER oligonucleotide may be referred to as the upstream oligonucleotide or probe in the complex in that, with the exception of its overlapping portion, it is largely 5' of a signal probe within the complex. On the target strand, the hybridization site of an INVADER oligonucleotide is downstream of the hybridization site of a signal probe.

The term "invasive cleavage structure" refers to a cleavage structure formed by the hybridization of two probes to a target nucleic acid, wherein the upstream probe overlaps the downstream probe by at least one base.

The term "DNA polymerase" as used herein refers to a protein that is encoded by a gene that is derived from the gene for a naturally occurring enzyme, said naturally occurring enzyme having nucleic acid synthetic activity. In some embodiments, the gene for the DNA

polymerase is altered relative to the gene for the naturally occurring enzyme, such that the DNA polymerase lacks or has reduced synthetic activity.

The term "synthetic activity" refers to the ability of an enzyme to catalyze DNA synthesis by addition of deoxyribonucleotide or deoxyribonucleotide analog units to a DNA chain using DNA or RNA as a template.

The term "target-complementary" refers to feature of a nucleic acid or a portion of an nucleic acid, said feature being that it is complementary to a target nucleic acid.

The terms "RNA-dependent" and "DNA-dependent", when used with reference to an enzymatic activity, refer to activity that occurs in response to the presence of an RNA or DNA target strand, respectively. A single enzyme may possess both types of activities and the presence of one type of activity does not imply the absence of any other activity. For example, description of an enzyme as having RNA-dependent 5' nuclease activity is not meant to indicate the absence of a DNA-dependent 5' nuclease activity, or any other activity. Similarly, an enzyme having a DNA-dependent activity may also have RNA-dependent and other activities.

The terms "target" and "template" are used interchangeably, and refer to a nucleic acid to be detected or analyzed. In some embodiments, the target is a nucleic acid to which one or more oligonucleotides or probes are hybridized. In a cleavage embodiment, one or more oligonucleotides may hybridize to form a cleavage structure, the formation of which may be used for detection or analysis of said target nucleic acid. The target nucleic acids include but are not limited to, single and double stranded DNA or RNA, modified nucleic acids (e.g., methylated nucleic acids), chimeric nucleic acids, peptide nucleic acids and the like.

"Nucleic acid molecule" refers to any nucleic acid containing molecule. The term encompasses sequences that include base analogs of DNA and RNA including, but not limited to, 4-acetylcytosine, 8-hydroxy-N6-methyladenosine, aziridinylcytosine, pseudoisocytosine, 5-(carboxyhydroxymethyl) uracil, 5-fluorouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxymethylaminomethyluracil, dihydrouracil, inosine, N6-isopentenyladenine, 1-methyladenine, 1-methylpseudouracil, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-methyladenine, 7-methylguanine, 5-methylaminomethyluracil,

5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine,
5'-methoxycarbonylmethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine,
uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, oxybutoxosine, pseudouracil,
queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil,
5 N-uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, pseudouracil, queosine,
2-thiocytosine, and 2,6-diaminopurine.

The term "overlap" as used herein in reference to cleavage structures, refers to a
feature of an invasive cleavage structure, said feature being that the 3' end of an upstream
INVADER oligonucleotide overlaps with the hybridized region of the signal probe by at least
one base. An overlap may be created by a duplication of sequence between the 3' portion of
the upstream INVADER oligonucleotide and the 5' portion of the target-complementary
region of the downstream probe oligonucleotide. The region of sequence so duplicated may
be as small as a single base. Regardless of the length of the duplicated sequence (*i.e.*, the
overlap) the 3' terminal base of the upstream INVADER oligonucleotide need not be
complementary to the target strand, and may be any nucleotide. In some embodiments, this
terminal nucleotide may be replaced by a moiety having chemical features similar to a
nucleotide such as a nucleotide analog or an aromatic ring compound. Indeed, any chemical
moiety that provides an overlap and promotes efficient invasive cleavage is contemplated. In
an alternative embodiment, the overlap need not involve any duplication of sequence between
the target-complementary regions of the two probes (Lyamichev *et al.*, Nat. Biotechnol.,
17:292 [1999], U.S. Patent 5,985,557). In this embodiment, the INVADER and signal probes
have regions complementary to adjacent regions of the target that are contiguous and that do
not overlap. When no sequence is shared, the 3' end of the upstream INVADER
oligonucleotide includes at least one additional nucleotide or nucleotide-like analog that is not
complementary to the target strand. This can be referred to as a physical overlap, in contrast
to a sequence overlap. An overlap of either type will satisfy the requirement for overlap of
the invasive cleavage of the INVADER assay.

As used herein, the terms "N-terminal" and "C-terminal" in reference to polypeptide
sequences refer to regions of polypeptides including portions of the N-terminal and C-terminal
regions of the polypeptide, respectively. A sequence that includes a portion of the N-terminal

region of polypeptide includes amino acids predominantly from the N-terminal half of the polypeptide chain, but is not limited to such sequences. For example, an N-terminal sequence may include an interior portion of the polypeptide sequence including bases from both the N-terminal and C-terminal halves of the polypeptide. The same applies to C-terminal regions.

One example of such a sequence is shown in Figure 6, wherein the "polymerase domain" may be referred to a C-terminal region of the polypeptide even though it contains amino acids that are part of the N-terminal half of the polypeptide. N-terminal and C-terminal regions may, but need not, include the amino acid defining the ultimate N-terminal and C-terminal ends of the polypeptide, respectively.

As used herein, the terms "complementary" or "complementarity" are used in reference to polynucleotides (*i.e.*, a sequence of nucleotides such as an oligonucleotide or a target nucleic acid) related by the base-pairing rules. For example, for the sequence "A-G-T," is complementary to the sequence "T-C-A." Complementarity may be "partial," in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there may be "complete" or "total" complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as detection methods which depend upon binding between nucleic acids.

The term "homology" refers to a degree of identity. There may be partial homology or complete homology. A partially identical sequence is one that is less than 100% identical to another sequence.

As used herein, the term "hybridization" is used in reference to the pairing of complementary nucleic acids. Hybridization and the strength of hybridization (*i.e.*, the strength of the association between the nucleic acids) is impacted by such factors as the degree of complementarity between the nucleic acids, stringency of the conditions involved, the T_m of the formed hybrid, and the G:C ratio within the nucleic acids.

As used herein, the term " T_m " is used in reference to the "melting temperature." The melting temperature is the temperature at which a population of double-stranded nucleic acid molecules becomes half dissociated into single strands. The equation for calculating the T_m of

nucleic acids is well known in the art. As indicated by standard references, a simple estimate of the T_m value may be calculated by the equation: $T_m = 81.5 + 0.41(\% G + C)$, when a nucleic acid is in aqueous solution at 1 M NaCl (*See e.g.*, Anderson and Young, Quantitative Filter Hybridization, in *Nucleic Acid Hybridization* (1985). Other references include more sophisticated computations which take structural as well as sequence characteristics into account for the calculation of T_m .

As used herein the term "stringency" is used in reference to the conditions of temperature, ionic strength, and the presence of other compounds, under which nucleic acid hybridizations are conducted. With "high stringency" conditions, nucleic acid base pairing will occur only between nucleic acid fragments that have a high frequency of complementary base sequences. Thus, conditions of "weak" or "low" stringency are often required when it is desired that nucleic acids which are not completely complementary to one another be hybridized or annealed together.

The term "gene" refers to a DNA sequence that comprises control and coding sequences necessary for the production of a polypeptide or precursor. The polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired enzymatic activity is retained.

The term "wild-type" refers to a gene or gene product which has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designed the "normal" or "wild-type" form of the gene. In contrast, the term "modified" or "mutant" refers to a gene or gene product which displays modifications in sequence and or functional properties (*i.e.*, altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are identified by the fact that they have altered characteristics when compared to the wild-type gene or gene product.

The term "recombinant DNA vector" as used herein refers to DNA sequences containing a desired coding sequence and appropriate DNA sequences necessary for the expression of the operably linked coding sequence in a particular host organism. DNA sequences necessary for expression in prokaryotes include a promoter, optionally an operator sequence, a ribosome binding site and possibly other sequences. Eukaryotic cells are known

to utilize promoters, polyadenylation signals and enhancers.

The term "oligonucleotide" as used herein is defined as a molecule comprising two or more deoxyribonucleotides or ribonucleotides, preferably at least 5 nucleotides, more preferably at least about 10-15 nucleotides and more preferably at least about 15 to 30 nucleotides. The exact size will depend on many factors, which in turn depends on the ultimate function or use of the oligonucleotide. The oligonucleotide may be generated in any manner, including chemical synthesis, DNA replication, reverse transcription, or a combination thereof.

Because mononucleotides are reacted to make oligonucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage, an end of an oligonucleotide is referred to as the "5' end" if its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose ring and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of a subsequent mononucleotide pentose ring. As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide, also may be said to have 5' and 3' ends. A first region along a nucleic acid strand is said to be upstream of another region if the 3' end of the first region is before the 5' end of the second region when moving along a strand of nucleic acid in a 5' to 3' direction.

When two different, non-overlapping oligonucleotides anneal to different regions of the same linear complementary nucleic acid sequence, and the 3' end of one oligonucleotide points towards the 5' end of the other, the former may be called the "upstream" oligonucleotide and the latter the "downstream" oligonucleotide.

The term "primer" refers to an oligonucleotide which is capable of acting as a point of initiation of synthesis when placed under conditions in which primer extension is initiated.

An oligonucleotide "primer" may occur naturally, as in a purified restriction digest or may be produced synthetically.

A primer is selected to be "substantially" complementary to a strand of specific sequence of the template. A primer must be sufficiently complementary to hybridize with a template strand for primer elongation to occur. A primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment may be

attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the strand. Non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the template to hybridize and thereby form a template primer complex for synthesis of the extension product of the primer.

"Hybridization" methods involve the annealing of a complementary sequence to the target nucleic acid (the sequence to be detected; the detection of this sequence may be by either direct or indirect means). The ability of two polymers of nucleic acid containing complementary sequences to find each other and anneal through base pairing interaction is a well-recognized phenomenon. The initial observations of the "hybridization" process by Marmur and Lane, *Proc. Natl. Acad. Sci. USA* 46:453 (1960) and Doty *et al.*, *Proc. Natl. Acad. Sci. USA* 46:461 (1960) have been followed by the refinement of this process into an essential tool of modern biology.

With regard to complementarity, it is important for some diagnostic applications to determine whether the hybridization represents complete or partial complementarity. For example, where it is desired to detect simply the presence or absence of pathogen DNA (such as from a virus, bacterium, fungi, mycoplasma, protozoan) it is only important that the hybridization method ensures hybridization when the relevant sequence is present; conditions can be selected where both partially complementary probes and completely complementary probes will hybridize. Other diagnostic applications, however, may require that the hybridization method distinguish between partial and complete complementarity. It may be of interest to detect genetic polymorphisms. For example, human hemoglobin is composed, in part, of four polypeptide chains. Two of these chains are identical chains of 141 amino acids (alpha chains) and two of these chains are identical chains of 146 amino acids (beta chains). The gene encoding the beta chain is known to exhibit polymorphism. The normal allele encodes a beta chain having glutamic acid at the sixth position. The mutant allele encodes a beta chain having valine at the sixth position. This difference in amino acids has a profound (most profound when the individual is homozygous for the mutant allele) physiological impact known clinically as sickle cell anemia. It is well known that the genetic basis of the amino acid change involves a single base difference between the normal allele DNA sequence and

the mutant allele DNA sequence.

The complement of a nucleic acid sequence as used herein refers to an oligonucleotide which, when aligned with the nucleic acid sequence such that the 5' end of one sequence is paired with the 3' end of the other, is in "antiparallel association." Certain bases not commonly found in natural nucleic acids may be included in the nucleic acids of the present invention and include, for example, inosine and 7-deazaguanine. Complementarity need not be perfect; stable duplexes may contain mismatched base pairs or unmatched bases. Those skilled in the art of nucleic acid technology can determine duplex stability empirically considering a number of variables including, for example, the length of the oligonucleotide, base composition and sequence of the oligonucleotide, ionic strength and incidence of mismatched base pairs.

The term "label" as used herein refers to any atom or molecule which can be used to provide a detectable (preferably quantifiable) signal, and which can be attached to a nucleic acid or protein. Labels may provide signals detectable by fluorescence, radioactivity, colorimetry, gravimetry, X-ray diffraction or absorption, magnetism, enzymatic activity, and the like. A label may be a charged moiety (positive or negative charge) or alternatively, may be charge neutral.

The term "cleavage structure" as used herein, refers to a structure which is formed by the interaction of at least one probe oligonucleotide and a target nucleic acid to form a complex having at least one region of base pairing before the probe and target, the resulting structure being cleavable by a cleavage agent, including but not limited to an enzyme. The cleavage structure is a substrate for specific cleavage by the cleavage agent, in contrast to a nucleic acid molecule which is a substrate for non-specific cleavage by agents such as phosphodiesterases which cleave nucleic acid molecules without regard to secondary structure (*i.e.*, no formation of a duplexed structure is required).

The terms "cleavage means" and "cleavage agent" as used herein refer to any agent which is capable of cleaving a cleavage structure, including but not limited to enzymes (*e.g.*, polymerases and flap endonucleases). The cleavage means may include native polymerases having 5' nuclease activity (*e.g.*, *Taq* DNA polymerase, *E. coli* DNA polymerase I) and, more specifically, modified polymerases having 5' nuclease but lacking synthetic activity.

"Structure-specific nucleases" or "structure-specific enzymes" are enzymes which recognize specific secondary structures in a nucleic molecule and cleave these structures. The cleavage means of the invention cleave a nucleic acid molecule in response to the formation of cleavage structures; it is not necessary that the cleavage means cleave the cleavage structure at any particular location within the cleavage structure.

The cleavage means is not restricted to enzymes having solely 5' nuclease activity. The cleavage means may include nuclease activity provided from a variety of sources including, but not limited to, the CLEAVASE enzymes, the FEN-1 endonucleases (including RAD2 and XPG proteins), *Taq* DNA polymerase and *E. coli* DNA polymerase I.

The term "thermostable" when used in reference to an enzyme, such as a 5' nuclease, indicates that the enzyme is functional or active (*i.e.*, can perform catalysis) at an elevated temperature, *e.g.*, at about 55°C or higher.

The term "cleavage products" as used herein, refers to products generated by the reaction of a cleavage means with a cleavage structure (*i.e.*, the treatment of a cleavage structure with a cleavage means).

The term "non-target cleavage product" refers to a product of a cleavage reaction which is not derived from the target nucleic acid. As discussed above, in the methods of the present invention, cleavage of the cleavage structure occurs within the probe oligonucleotide. The fragments of the probe oligonucleotide generated by this target nucleic acid-dependent cleavage are "non-target cleavage products."

The term "substantially single-stranded" when used in reference to a nucleic acid substrate means that the substrate molecule exists primarily as a single strand of nucleic acid in contrast to a double-stranded substrate which exists as two strands of nucleic acid which are held together by inter-strand base pairing interactions.

The term "sequence variation" as used herein refers to differences in nucleic acid sequence between two nucleic acids. For example, a wild-type structural gene and a mutant form of this wild-type structural gene may vary in sequence by the presence of single base substitutions and/or deletions or insertions of one or more nucleotides. These two forms of the structural gene are said to vary in sequence from one another. A second mutant form of the structural gene may exist. This second mutant form is said to vary in sequence from both

the wild-type gene and the first mutant form of the gene.

The term "liberating" as used herein refers to the release of a nucleic acid fragment from a larger nucleic acid fragment, such as an oligonucleotide, by the action of a 5' nuclease such that the released fragment is no longer covalently attached to the remainder of the oligonucleotide.

The term " K_m " as used herein refers to the Michaelis-Menten constant for an enzyme and is defined as the concentration of the specific substrate at which a given enzyme yields one-half its maximum velocity in an enzyme catalyzed reaction.

The term "nucleotide analog" as used herein refers to modified or non-naturally occurring nucleotides such as 7-deaza purines (*i.e.*, 7-deaza-dATP and 7-deaza-dGTP). Nucleotide analogs include base analogs and comprise modified forms of deoxyribonucleotides as well as ribonucleotides.

The term "polymorphic locus" is a locus present in a population which shows variation between members of the population (*i.e.*, the most common allele has a frequency of less than 0.95). In contrast, a "monomorphic locus" is a genetic locus at little or no variations seen between members of the population (generally taken to be a locus at which the most common allele exceeds a frequency of 0.95 in the gene pool of the population).

The term "microorganism" as used herein means an organism too small to be observed with the unaided eye and includes, but is not limited to bacteria, virus, protozoans, fungi, and ciliates.

The term "microbial gene sequences" refers to gene sequences derived from a microorganism.

The term "bacteria" refers to any bacterial species including eubacterial and archaeobacterial species.

The term "virus" refers to obligate, ultramicroscopic, intracellular parasites incapable of autonomous replication (*i.e.*, replication requires the use of the host cell's machinery).

The term "sample" in the present specification and claims is used in its broadest sense. On the one hand it is meant to include a specimen or culture (*e.g.*, microbiological cultures or cultured eukaryotic tissue cells). On the other hand, it is meant to include both biological and environmental samples.

Biological samples may be animal, including human, fluid, solid (*e.g.*, stool) or tissue, as well as liquid and solid food and feed products and ingredients such as dairy items, vegetables, meat and meat by-products, and waste. Biological samples may be obtained from all of the various families of domestic animals, as well as feral or wild animals, including, but not limited to, such animals as ungulates, bear, fish, lagamorphs, rodents, etc.

Environmental samples include environmental material such as surface matter, soil, water and industrial samples, as well as samples obtained from food and dairy processing instruments, apparatus, equipment, utensils, disposable and non-disposable items. These examples are not to be construed as limiting the sample types applicable to the present invention.

The term "source of target nucleic acid" refers to any sample which contains nucleic acids (RNA or DNA). Particularly preferred sources of target nucleic acids are biological samples including, but not limited to cultured cells, blood, saliva, cerebral spinal fluid, pleural fluid, milk, lymph, sputum, and semen.

An oligonucleotide is said to be present in "excess" relative to another oligonucleotide (or target nucleic acid sequence) if that oligonucleotide is present at a higher molar concentration relative to that of the other oligonucleotide (or target nucleic acid sequence). Typically, when present in excess, the probe oligonucleotide will be present at least a 100-fold molar excess.

A sample "suspected of containing" a first and a second target nucleic acid may contain either, both or neither target nucleic acid molecule.

The terms "polymerization means" or "polymerization agent" refer to any agent capable of facilitating the addition of nucleoside triphosphates to an oligonucleotide.

The terms "ligation means" or "ligation agent" refer to any agent capable of facilitating the ligation (*i.e.*, the formation of a phosphodiester bond between a 3'-OH and a 5'-P located at the termini of two strands of nucleic acid). Preferred ligation means comprise DNA ligases and RNA ligases.

The term "reactant" is used herein in its broadest sense. In some embodiments, the reactant can comprise an enzymatic reactant, a chemical reactant or ultraviolet light (ultraviolet light, particularly short wavelength ultraviolet light is known to break

oligonucleotide chains). Any agent capable of reacting with an oligonucleotide to either shorten (*i.e.*, cleave) or elongate the oligonucleotide is encompassed within the term "reactant."

The terms "adduct" and "moiety" are used interchangeably, are used herein in their
5 broadest sense to indicate any compound or element which can be added to an
oligonucleotide. An adduct may be charged (positively or negatively) or may be charge
neutral. An adduct may be added to the oligonucleotide via covalent or non-covalent
linkages. Examples of adducts, include but are not limited to indodicarbocyanine dye
amidites, amino-substituted nucleotides, ethidium bromide, ethidium homodimer,
10 (1,3-propanediamino)propidium, (diethylenetriamino)propidium, thiazole orange,
(N-N'-tetramethyl-1,3-propanediamino)propyl thiazole orange,
(N-N'-tetramethyl-1,2-ethanediamino)propyl thiazole orange, thiazole orange-thiazole orange
homodimer (TOTO), thiazole orange-thiazole blue heterodimer (TOTAB), thiazole
orange-ethidium heterodimer 1 (TOED1), thiazole orange-ethidium heterodimer 2 (TOED2)
and fluorescein-ethidium heterodimer (FED), psoralens, biotin, streptavidin, avidin, etc.

The term "recombinant DNA molecule" as used herein refers to a DNA molecule that
comprises of segments of DNA joined together by means of molecular biological techniques.

The term "recombinant protein" or "recombinant polypeptide" as used herein refers to
a protein molecule that is expressed from a recombinant DNA molecule.

As used herein the term "portion" when in reference to a protein (as in "a portion of a
given protein") refers to fragments of that protein. The fragments may range in size from
four amino acid residues to the entire amino acid sequence minus one amino acid.

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide or
polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or
25 synthetic origin which may be single- or double-stranded, and represent the sense or antisense
strand. Similarly, "amino acid sequence" as used herein refers to peptide or protein sequence.

"Peptide nucleic acid" ("PNA") as used herein refers to a molecule which comprises an
oligomer to which an amino acid residue, such as lysine, and an amino group have been
added. These small molecules, also designated anti-gene agents, stop transcript elongation by
30 binding to their complementary strand of nucleic acid (Nielsen *et al.*, Anticancer Drug Des.

8:53 [1993]).

As used herein, the terms "purified" or "substantially purified" refer to molecules, either nucleic or amino acid sequences, that are removed from an environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they were associated in the starting environment. The starting environment may be a natural one, as in the isolation of a non-recombinant protein, or it may be a created environment, as is the isolation of a recombinant protein from a host cell. For example, recombinant CLEAVASE nucleases may be expressed in bacterial host cells and the nucleases may be purified by the removal of host cell proteins; the percent of these recombinant nucleases is thereby increased in the sample. An "isolated polynucleotide" or "isolated oligonucleotide" is therefore a substantially purified polynucleotide.

As used herein, the term "fusion protein" refers to a chimeric protein containing the protein of interest joined to an exogenous protein fragment. The fusion partner may enhance solubility of recombinant chimeric protein as expressed in a host cell, may provide an affinity tag (e.g., a his-tag) to allow purification of the recombinant fusion protein from the host cell or culture supernatant, or both. If desired, the fusion protein may be removed from the protein of interest by a variety of enzymatic or chemical means known to the art.

As used herein, the terms "chimeric protein" and "chimerical protein" refer to a single protein molecule that comprises amino acid sequence portions derived from two or more parent proteins. These parent molecules may be similar proteins from genetically distinct origins, different proteins from a single organism, or dissimilar proteins from different organisms. By way of example but not by way of limitation, a chimeric structure-specific nuclease of the present invention may contain a mixture of amino acid sequences that have been derived from DNA polymerase genes from two or more of the organisms having such genes, combined to form a non-naturally occurring 5' nuclease. The term "chimerical" as used herein is not intended to convey any particular proportion of contribution from the naturally occurring genes, nor limit the manner in which the portions are combined. Any chimeric structure-specific nuclease constructs having cleavage activity as determined by the testing methods described herein, for example, are improved cleavage agents within the scope of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The INVADER invasive cleavage reaction has been shown to be useful in the detection of RNA target strands (*See e.g.*, U.S. Patent 6,001,567, incorporated herein by reference in its entirety). As with the INVADER assay for the detection of DNA (Lyamichev *et al.*, Nat. Biotechnol., 17:292 [1999]), the reactions may be run under conditions that permit the cleavage of many copies of a probe for each copy of the target RNA present in the reaction. In one embodiment, the reaction may be performed at a temperature close to the melting temperature (T_m) of the probe that is cleaved, such that the cleaved and uncleaved probes readily cycle on and off the target strand without temperature cycling. Each time a full-length probe binds to the target in the presence of the INVADER oligonucleotide, it may be cleaved by a 5' nuclease enzyme, resulting in an accumulation of cleavage product. The accumulation is highly specific for the sequence being detected, and may be configured to be proportional to both time and target concentration of the reaction. In another embodiment, the temperature of the reaction may be shifted (*i.e.*, it may be raised to a temperature that will cause the probe to dissociate) then lowered to a temperature at which a new copy of the probe hybridizes to the target and is cleaved by the enzyme. In a further embodiment, the process of raising and lowering the temperature is repeated many times, or cycled, as it is in PCR (Mullis and Faloona, Methods in Enzymology, 155:335 [1987], Saiki *et al.*, Science 230:1350 [1985]).

As noted above, 5' nucleases of Pol A type DNA polymerases are preferred for cleavage of an invasive cleavage structure that comprises an RNA target strand. The present invention provides enzymes having improved performance in detection assays based on the cleavage of a structure comprising RNA. In particular, the altered polymerases of the present invention exhibit improved performance in detection assays based on the cleavage of a DNA member of an invasive cleavage structure that comprises an RNA target strand.

The 5' nucleases of the present invention may be derived from Pol A type DNA polymerases. The terminology used in describing the alterations made in this class of 5' nucleases relates to the descriptions of DNA polymerase structures known in the art. The Klenow fragment of the Pol A polymerase from *E. coli* (the C-terminal two thirds, which has

the DNA synthesizing activity but lacks the 5' nuclease activity) has been described as having a physical form resembling a right hand, having an open region called the "palm", and a cleft that holds the primer/template duplex defined on one side by a "fingers" domain and on the other by a "thumb" domain (Joyce and Steitz, Trends in Biochemical Science 12:288 [1987]).

5 This is shown schematically in Figure 5. Because this physical form has proved to be common to all Pol A DNA polymerases and to a number of additional template-dependent polymerizing enzymes such as reverse transcriptases, the hand terminology has become known in the art, and the sites of activity in these enzymes are often described by reference to their position on the hand. For reference, and not intended as a limitation on the present invention,
10 the palm is created from roughly the first 200 amino acids of the polymerase domain, the thumb from the middle 140, and the fingers by the next 160, with the base of the cleft formed from the remaining regions (Figures 6). Although some enzymes may deviate from these structural descriptions, the equivalent domains and sequences corresponding to such domains may be identified by sequence homology to known enzyme sequences, by comparison of
15 enzyme crystal structures, and other like methods.

In creating the improved enzymes of the present invention, several approaches have been taken, although the present invention is not limited to any particular approach. First two DNA polymerases, Taq and Tth, that have different rates of DNA strand cleavage activity on RNA targets were compared. To identify domains related to the differences in activity, a
20 series of chimerical constructs was created and the activities were measured. This process identified two regions of the Tth polymerase that could, if transferred into the Taq polymerase, confer on the TaqPol an RNA-dependent cleavage activity equivalent to that of the native Tth protein. Once these regions were identified, the particular amino acids involved in the activity were examined. Since the two proteins are about 87 percent identical
25 in amino acid sequence overall, the identified regions had only a small number of amino acid differences. By altering these amino acids singly and in combinations, a pair of amino acids were identified in TthPol that, if introduced into the TaqPol protein, increased the rate of cleavage up to that of the native TthPol.

These data demonstrate two important aspects of the present invention. First, specific
30 amino acids can be changed to confer TthPol-like RNA-dependent cleavage activity on a

polymerase having a lesser activity. More broadly, however, these results provide regions of these polymerases that are involved in the recognition of the RNA-containing cleavage structure. Identification of these important regions, combined with published information on the relationships of other amino acids to the various functions of these DNA polymerases and computer-assisted molecular modeling during the development of the present invention have allowed a rational design approach to create additional improved 5' nucleases. The information also allowed a focused random mutagenesis approach coupled with a rapid screening procedure to quickly create and identify enzymes having improved properties. Using these methods of the present invention, a wide array of improved polymerases are provided.

The methods used in creating and selecting the improved 5' nucleases of the present invention are described in detail below and in the experimental examples. A general procedure for screening and characterizing the cleavage activity of any 5' nuclease is included in the experimental examples. The methods discussions are divided into the following sections: I) Creation and selection of chimerical constructs; II) Site-specific mutagenesis based on information from chimerical constructs; III) Site-specific mutagenesis based on molecular modeling and published physical studies; and IV) focused random mutagenesis.

I) Creation and selection of chimerical constructs

The PolA-type DNA polymerases, including but not limited to DNA polymerase enzymes from *Thermus* species, comprise two distinctive domains, the 5' nuclease and the polymerase domains, shown schematically in Figure 6. The polymerase domains reside in the C-terminal two-thirds of the proteins and are responsible for both DNA-dependent and RNA-dependent DNA polymerase activities. The N-terminal one-third portions contain the 5' nuclease domains. In the genus *Thermus* Pol A polymerase, the palm region consists of, roughly, amino acids 300-500, the thumb region includes amino acids 500-650, while the fingers region is formed by the remaining amino acids from 650 to 830 (Figure 6).

The derivatives, Taq DN RX HT and Tth DN RX HT, of Taq and TthPol used in many of the experiments of the present invention, and described herein, are modified to reduce synthetic activity and to facilitate chimera construction, but have 5' nuclease activity

essentially identical to unmodified TaqPol and TthPol. Unless otherwise specified, the TaqPol and TthPol enzymes of the following discussion refer to the DN RX HT derivative.

TthPol has a 4-fold higher cleavage rate with the IL-6 RNA template (shown in Figure 10) than TaqPol (shown in Figures 11 and 12), although the Taq and TthPols show similarities of cleavage in DNA target structures (Figure 10). Since the amino acid sequences of TaqPol and TthPol (Figures 8 and 9) share about 87% identity and greater than 92% similarity, the high degree of homology between the enzymes allowed creation of a series of chimeric enzymes between TthPol and TaqPol. The activity of the chimeric enzymes was used as a parameter to identify the region(s) of these proteins affecting RNA dependent 5' nuclease activity.

The chimeric constructs between TthPol and TaqPol genes shown schematically in Figures 7 and 19 were created by swapping DNA fragments defined by the restriction endonuclease sites, EcoRI and BamHI, common for both genes, the cloning vector site SalI and the new sites, NotI, BstBI and NdeI, created at the homologous positions of both genes by site directed mutagenesis. The restriction enzymes have been abbreviated as follows: EcoRI is E; NotI is N; BstBI is Bs; NdeI is D, BamHI is B, and SalI is S.

The activity of each chimeric enzyme was evaluated using the invasive signal amplification assay with the IL-6 RNA target (Figure 10), and the cycling cleavage rates shown in Figure 12 were determined as described in the Experimental Examples. Comparison of the cleavage rates of the first two chimeras, TaqTth(N) and TthTaq(N), created by swapping the polymerase and 5' nuclease domains at the NotI site (Figure 7), shows that TaqTth(N) has the same activity as TthPol, whereas its counterpart TthTaq(N) retains the activity of TaqPol (Figure 12). This result indicates that the higher cleavage rate of TthPol is associated with its polymerase domain and suggests an important role of the polymerase domain in the 5' nuclease activity.

The next step was to identify a minimal region of TthPol polymerase that would give rise to the TthPol-like RNA dependent 5' nuclease activity when substituted for the corresponding region of the TaqPol sequence. To this end, the TaqTth(N) chimera was selected to generate a series of new constructs by replacing its TthPol sequence with homologous regions of TaqPol. First, the N-terminal and C-terminal parts of the TaqPol

polymerase domain were substituted for the corresponding regions of TaqTth(N) using the common BamHI site as a breaking point to create TaqTth(N-B) and TaqTth(B-S) chimeras, respectively (Figure 7). TaqTth(N-B) which has the TthPol sequence between amino acids 328 and 593, is approximately 3 times more active than the TaqTth(B-S) and 40% more active than TthPol (Figure 12). This result establishes that the NotI-BamHI portion of the TthPol polymerase domain determines superior RNA-dependent 5' nuclease activity of TthPol.

From these data it was determined that a central portion of the TthPol, when used to replace the homologous portion of TaqPol (TaqTth(N-B) construct) could confer superior RNA recognition on the chimerical protein composed primarily of Taq protein. In fact, the cycling rate of this chimerical protein exceeded that of the parent TthPol. Comparison of chimeras that included sub-portions of the activity-improving region of TthPol, approximately 50% of the region in each case (*See*, TaqTth(N-D) and TaqTth(D-B), Figures 7 and 12) showed no significant improvement in RNA dependent activity as compared to the parent TaqPol. This result indicates that aspects of each half of the region are required for this activity. A construct having an only slightly smaller portion of the Tth insert portion (TaqTth(Bs-B)) showed activity that was close to that of the parent TthPol protein, but which was less than that of the TaqTth(N-B) construct.

II) Site-specific mutagenesis based on information from chimerical constructs

Comparison of the TthPol and TaqPol amino acid sequences between the BstBI and BamHI sites reveals only 25 differences (Figure 13). Among those, there are 12 conservative changes and 13 substitutions resulting in a change in charge. Since the analysis of the chimeric enzymes has suggested that some critical amino acid changes are located in both BstBI-NdeI and NdeI-BamHI regions of TthPol, site directed mutagenesis was used to introduce the TthPol specific amino acids into the BstBI-NdeI and NdeI-BamHI regions of the TaqTth(D-B) and TaqTth (N-D) chimeras, respectively. Six TthPol-specific substitutions were generated in the BstBI-NdeI region of the TaqTth(D-B) by single or double amino acid mutagenesis and only one double mutation, W417L/G418K, was able to restore the TthPol activity with the IL-6 RNA target (*See e.g.*, Figure 14). Similarly, 12 TthPol specific amino

acids were introduced at the homologous positions of the NdeI-BamHI region of the TaqTth(N-D) and only one of them, E507Q, increased the cleavage rate to the TthPol level (See e.g., Figure 14).

To confirm that the W417L, G418K and E507Q substitutions are sufficient to increase the TaqPol activity to the TthPol level, TaqPol variants carrying these mutations were created and their cleavage rates with the IL-6 RNA substrate were compared with that of TthPol. Figure 15 shows that the TaqPol W417L/G418K/E507Q and TaqPol G418K/E507Q mutants have 1.4 times higher activity than TthPol and more than 4 fold higher activity than TaqPol, whereas the TaqPol W417L/E507Q mutant has the same activity as TthPol, which is about 3 fold higher than TaqPol. These results demonstrate that K418 and Q507 of TthPol are important amino acids in defining its superior RNA dependent 5' nuclease activity compared to TaqPol.

The ability of these amino acids to improve the RNA dependent 5' nuclease activity of a DNA polymerase was tested by introducing the corresponding mutations into the polymerase A genes of two additional organisms: *Thermus filiformis* and *Thermus scotoductus*. TaqPol showed improved RNA dependent activity when it was modified to contain the W417L and E507Q mutations, which made it more similar at these residues to the corresponding residues of TthPol (K418 and Q507). The TfiPol was modified to have P420K and E507Q, creating TfiDN 2M, while the TscPol was modified to have E416K and E505Q, to create TscDN 2M. The activity of these enzymes for cleaving various DNA and RNA containing structures was determined as described in Example 1, using the IdT2, IrT3, hairpin and X-structures diagrammed in Figures 21 and 22, with the results shown in both Figure 25 and Table 7. Both enzymes have much less RNA-dependent cleavage activity than either the TthPol or the Taq 2M enzymes. However, introduction of the mutations cited above into these polymerases increased the RNA dependent cleavage activity over 2 fold compared to the unmodified enzymes (Figure 25). These results demonstrate that transferability of improved RNA dependent cleavage activity into a wide range of polymerases using the methods of the present invention.

III) Site-specific mutagenesis based on molecular modeling and published physical studies

The positions of the G418H and E507Q mutations in the crystal structure of a complex of the large fragment of TaqPol (Klentaq1) with a primer/template DNA determined by Li *et al.* (Li *et al.*, Protein Sci., 7:1116 [1998]) are shown in Figure 17. The E507Q mutation is located at the tip of the thumb subdomain at a nearest distance of 3.8 Å and 18 Å from the backbone phosphates of the primer and template strands, respectively. The interaction between the thumb and the minor groove of the DNA primer/template was previously suggested by the co-crystal structures of Klenow fragment DNA polymerase I (Breese *et al.*, Science 260:352 [1993]) and TaqPol (Eom *et al.*, Nature 382:278 [1996]). Deletion of a 24 amino acid portion of the tip of the thumb in Klenow fragment, corresponding to amino acids 494-518 of TaqPol, reduces the DNA binding affinity by more than 100-fold (Minnick *et al.*, J. Biol. Chem., 271:24954 [1996]). These observations are consistent with the hypothesis that the thumb region, which includes the E507 residue, is involved in interactions with the upstream substrate duplex.

The W417L and the G418K mutations in the palm region of TaqPol (Figure 17) are located approximately 25 Å from the nearest phosphates of the template and upstream strands, according to the co-crystal structures of TaqPol with duplex DNA bound in the polymerizing mode (Li *et al.*, Protein Sci., 7:1116 [1998], Eom *et al.*, Nature 382:278 [1996]). The same distance was observed between the analogous W513 and P514 amino acids of Klenow fragment and the template strand of DNA bound in the editing mode (Breese *et al.*, Science 260:352 [1993]). Thus, no interactions between TaqPol and the overlapping substrate can be suggested from the available co-crystal studies for this region.

Although an understanding of the mechanism of action of the enzymes is not necessary for the practice of the present invention and the present invention is not limited to any mechanism of action, it is proposed that the amino acids at positions 417 and 418 in the palm region of TaqPol interact with the upstream substrate duplex only when the enzyme functions as a 5' nuclease, but no interaction with these amino acids occurs when TaqPol switches into polymerizing mode. This hypothesis suggests a novel mode of substrate binding by DNA

polymerases called here the “5’ nuclease mode.” Several lines of evidence support this hypothesis. The study of the chimeric enzymes described here clearly separates regions of the polymerase domain involved in the 5’ nuclease and polymerase activities. Accordingly, the W417L and G418K mutations, together with the E507Q mutation, affect the 5’ nuclease activity of TaqPol on substrates having an RNA target strand (Figure 15), but have no effect on either RNA-dependent or DNA-dependent DNA polymerase activities (Figure 16). On the other hand, mutations in the active site of TaqPol, such as R573A, R587A, E615A, R746A, N750A and D785N, which correspond to substitutions in Klenow fragment of *E.coli* DNA Pol I that affect both polymerase activity and substrate binding affinity in the polymerizing mode (Polesky *et al.*, J. Biol. Chem., 265:14579 [1990], Polesky *et al.*, J. Biol. Chem., 267:8417 [1992], Pandey *et al.*, Eur. J. Biochem., 214:59 [1993]) were shown to have little or no effect on the 5’ nuclease activity. Superposition of the polymerase domains of TaqPol (Eom *et al.*, Nature 382:278 [1996]), *E.coli* Pol I and *Bacillus stearothermophilus* Pol I (Kiefer *et al.*, Nature 391:304 [1998]) using the programs DALI (Holm and Sander, J. Mol. Biol., 233:123 [1993], Holm and Sander, Science 273:595 [1996]) and Insight II (Molecular Simulation Inc., Naperville, IL) shows that the palm region of TaqPol between amino acids 402-451, including W417 and G418, is structurally highly conserved between the three polymerases, although there is no structural similarity between the rest of the palm subdomains. This observation suggests an important role for this region in eubacterial DNA polymerases.

The 5’ nuclease and polymerase activities should be precisely synchronized to create a nicked structure rather than a gap or an overhang that could cause a deletion or an insertion during Okazaki fragment processing or DNA repair, if ligase joins the ends inappropriately. According to the previously proposed model (Kaiser *et al.*, J. Biol. Chem., 274:21387 [1999]), the 3’ terminal nucleotide of the upstream strand is sequestered by the 5’ nuclease domain to prevent its extension, thus halting synthesis. The interaction with the 3’ nucleotide apparently activates the 5’ nuclease that endonucleolitically removes the displaced 5’ arm of the downstream strand. This cleavage occurs by the precise incision at the site defined by the 3’ nucleotide, thus creating the nick. This model requires a substantial rearrangement of the substrate-enzyme complex, which may include a translocation of the complex to the 5’

nuclease mode to separate the primer/template from the polymerase active site.

It is possible that a relocation of the substrate away from the polymerase active site could be induced by the interaction between the duplex formed between the template and incoming strands and the crevice formed by the finger and thumb subdomains. Such an interaction could force conformational transitions in the thumb that would bring the template/primer duplex into close contact with the W417 and G418 amino acids. Significant flexibility of the thumb has been previously reported that might explain such changes (Beese *et al.*, Science 260:352 [1993], Eom *et al.*, Nature 382:278 [1996], Ollis *et al.*, Nature 313:762 [1985], Kim *et al.*, Nature 376:612 [1995], Korolev *et al.*, Proc. Natl. Acad. Sci., 92:9264 [1995], Li *et al.*, EMBO J., 17:7514 [1998]). Additional conformational changes of the fingers domain that might help to open the crevice, such as the transition from the 'closed' to the 'open' structure described by Li *et al.* (Li *et al.*, EMBO J., 17:7514 [1998]), are consistent with this model. It may be that the 5' nuclease binding mode was not observed in any of the published co-crystal structures of a DNA Pol I because the majority of the structures were solved for the polymerase domain only, with a template/primer substrate rather than with an overlapping 5' nuclease substrate.

K_m values of 200-300 nM have been determined for TaqPol, TthPol and TaqPol G418K/E507Q for the RNA containing substrate. These values are much higher than the K_m value of <1 nM estimated for TthPol with an all-DNA overlapping substrate suggesting that the RNA template adversely affects substrate binding. The low affinity could be explained by the unfavorable interaction between the enzyme and either the A-form duplex adopted by the substrate with an RNA target, or the ribose 2' hydroxyls of the RNA strand. Between these two factors, the latter seems more likely, since the 5' nucleases of eubacterial DNA polymerases can efficiently cleave substrates with an RNA downstream probe (Lyamichev *et al.*, Science 260:778 [1993]), which would presumably have an A-form. Further, the co-crystal studies suggest that the template/primer duplex partially adopts a conformation close to A-form in its complex with DNA polymerase (Eom *et al.*, Nature 382:278 [1996], Kiefer *et al.*, Nature 391:304 [1998], Li *et al.*, EMBO J., 17:7514 [1998]). The G418K/E507Q mutations increase the k_{cat} of TaqPol more than two fold, but have little effect

on K_m . Such an effect would be expected if the mutations position the substrate in an orientation more appropriate for cleavage rather than simply increasing the binding constant.

In addition to the mutational analysis described above, another approach to studying specific regions of enzymes, enzyme structure-function relationships, and enzyme-substrate interaction is to investigate the actual, physical structure of the molecule.

With the advances in crystallographic, NMR, and computer and software technology, study of molecular structure has become a viable tool for those interested in the configuration, organization, and dynamics of biomolecules. Molecular modeling has increased the understanding of the nature of the interactions that underlie the structure of proteins and how proteins interact functionally with substrate. Numerous publications describing the structures of various polymerases or polymerase protein portions, HIV reverse transcriptase, and other nucleic acid binding proteins have provided mechanistic insights into protein conformation, changes in conformation, and molecular interactions necessary for function.

As an example, the report by Doublie *et al.* (Doublie *et al.*, Nature 391:251 [1998]) discloses the crystal structure of T7 DNA polymerase and provides information about which amino acid regions are likely to have an affect on substrate binding, which are required to contact the substrate for polymerization, and which amino acids bind cofactors, such as metal ions. It is noted in this paper and others that many of the polymerases share not only sequence similarity, but structural homology as well. When certain structural domains of different polymerases are superimposed (for example, T7 polymerase, Klenow fragment editing complex, the unliganded Taq DNA polymerase and the Taq Polymerase-DNA complex) conserved motifs are clearly discernable.

Specifically, combining the information from all of these different structural sources and references, a model of the protein interacting with DNA, RNA, or heteroduplex can be made. The model can then be examined to identify amino acids that may be involved in substrate recognition or substrate contact. Changes in amino acids can be made based on these observations, and the effects on the various activities of the 5' nuclease proteins are assessed using screening methods such as those of the present invention, described in the experimental examples.

The domain swapping analysis discussed previously demonstrated that sequences of

TthDN that are important in RNA-dependent 5' nuclease activity lie in the polymerase domain of the protein. Therefore, study of structural data of the polymerase domain with respect to nucleic acid recognition provides one method of locating amino acids that, when altered, alter RNA recognition in a 5' nuclease reaction. For example, analysis conducted during the development of the present invention examined published analyses relating to primer/template binding by the polymerase domain of *E. coli* Pol 1, the Klenow fragment. Table 1 shows a sampling of kinetic constants determined for the Klenow fragment, and shows the effects a number of mutations on these measurements. The corresponding or similarly positioned amino acids in the TaqPol are indicated in the right hand column. It was postulated that mutations having a noticeable impact on the interactions of the Klenow fragment with the DNA template or the primer/template duplex, as indicated by changes in K_d and Relative DNA affinity values, might also have effects when made at the corresponding sites in TaqPol and related chimerical or mutant derivatives. A selection of the mutations that produced a higher K_d value or a lower Relative DNA affinity value when introduced into the Klenow fragment were created and examined in TaqPol. These Taq derivatives include, but are not limited to, those indicated by asterisks in the right hand column of Table 1.

For some Klenow variants, such as the R682 mutants, selection for testing was not made based on the DNA affinity measurements, but because molecular modeling suggested interaction between some aspect of the template/primer duplex and that amino acid. Similarly, additional regions of Taq polymerase (or Taq derivatives) were targeted for mutagenesis based on structural data and information from molecular modeling. Based on modeling, the thumb region was postulated to contact an RNA template. Thus, amino acids in the thumb region were looked for that, if altered, might alter that contact. For example, Figures 6 and 17 show that amino acids 502, 504, and 507 are located at the tip of the thumb. It was postulated that altering these amino acids might have an affect on the enzyme-substrate interaction. Using the activity screening methods described In Example 1, mutations that produced beneficial effects were identified. This approach was used to create a number of improved enzymes. For example, TaqPol position H784, corresponding to Klenow amino acid H881, is an amino acid in the fingers region and, as such, may be involved in primer/template substrate binding. When the H881 amino acid in the Klenow enzyme is

replaced by alanine, the change decreases the affinity of the enzyme for DNA to only 30 to 40% of the wild type level. An analogous substitution was tested in a TaqPol-derived enzyme. Starting with the Taq derivative W417L/G418K/E507Q, amino acid 784 was changed from Histidine (H) to Alanine (A) to yield the W417L/G418K/E507Q/H784A mutant, termed Taq 4M. This variant showed improved 5' nuclease activity on the RNA test IrT1 (Figure 24) test substrate (data in Table 2). Amino acid R587 is in the thumb region, and was selected for mutation based on its close proximity to the primer/template duplex in computer models. When an R587A mutation was added to the Taq 4M variant, the activity on the test IrT1 test substrate was still further improved. In addition, the reduction, relative to the 4M variant, in cleavage of the X structure shown in Figure 22 constitutes an additional improvement in this enzyme's function.

Not all amino acid changes that reduce DNA binding in the polymerization affect the 5' nuclease activity. For example, mutations E615A, R677A, affecting amino acid that are also in the thumb and fingers domains, respectively, have either adverse effect, or no effect on the 5' nuclease activities, respectively, as measured using the test substrates in Figures 21 and 22, and compared to the parent variants that lacked these changes. The R677A mutation was added to, and compared with the TaqSS variant, while the E615A mutation was added to and compared with the Taq 4M variant. The test methods described herein provide a convenient means of analyzing any variant for the alterations in the cleavage activity of both invasive and noninvasive substrates, for both DNA and RNA containing structures. Thus, the present invention provides methods for identifying all suitable improved enzymes.

Alterations that might increase the affinity of the enzymes for the nucleic acid targets were also examined. Many of the mutations described above were selected because they caused the Klenow fragment enzyme to have decreased affinity for DNA, with the goal of creating enzymes more accepting of structures containing non-DNA strands. In general, the native DNA polymerases show a lower affinity for RNA/DNA duplexes, compared to their affinity for DNA/DNA duplexes. During the development of the present invention, it was sought to increase the general affinity of the proteins of the present invention for a nucleic acid substrate without restoring or increasing any preference for structures having DNA rather than RNA target strands. The substitution of amino acids having different charges was

examined as a means of altering the interaction between the proteins and the nucleic acid substrates. For example, it was postulated that addition of positively charged amino acid residues, such as lysine (K), might increase the affinity of a protein for a negatively charged nucleic acid.

As noted above, alterations in the thumb region could affect the interactions of the protein with the nucleic acid substrate. In one example, the mutation G504K (tip of the thumb) was introduced in Taq4M and caused an enhancement of nuclease activity by 15% on an RNA target. Additional positively charged mutations (A502K and E507K) further improve the RNA target dependent activity by 50% compared to the parent Taq4M enzyme.

The use of data from published studies and molecular modeling, in combination with results accrued during the development of the present invention allowed the identification of regions of the proteins in which changes of amino acids would be likely to cause observable differences in at least one aspect of cleavage function. While regions could be targeted in this way, it was observed that changes in different amino acids, even if near or immediate neighbors in the protein, could have different effects. For example, while the A502K substitution created a marked increase in the RNA-dependent cleavage activity of Taq 4M, changing amino acid 499 from G to a K, only 3 amino acids away from 502, gave minimal improvement. As can be seen in the Experimental Examples, the approach of the present invention was to change several amino acids in a candidate region, either alone or in combination, then use the screening method provided in Example 1 to rapidly assess the effects of the changes. In this way, the rational design approach is easily applied to the task of protein engineering.

In addition to the thumb, palm, and hand regions found in the polymerase domain of these proteins, regions that are specific to 5' nucleases and nuclease domains were examined.

Comparative studies on a variety of 5' nucleases have shown that, though the amino acid sequences vary dramatically from enzyme to enzyme, there are structural features common to most. Two of these features are the helix-hairpin-helix motif (H-h-H) and the arch or loop structure. The H-h-H motif is believed to mediate non-sequence specific DNA binding. It has been found in at least 14 families of proteins, including nucleases, N-glycosylases, ligases, helicases, topoisomerases, and polymerases (Doherty *et al.*, Nucl. Acid. Res., 24:2488 [1996]).

The crystallographic structure of rat DNA polymerase pol β bound to a DNA template-primer shows non-specific hydrogen bonds between the backbone nitrogens of the pol β HhH motif and the phosphate oxygens of the primer of the DNA duplex (Pelletier *et al.*, Science 264:1891 [1994]). Because the HhH domain of 5' nuclease domains of Taq and Tth polymerases may function in a similar manner, it is contemplated that mutations in the HhH region of the enzyme alter activity. Mutations may be introduced to alter the shape and structure of the motif, or to change the charge of the motif to cause increased or decreased affinity for substrate.

Another structure common to many 5' nucleases from diverse sources such as eukaryotes, eubacteria, archaea and phage, is the arch or loop domain. The crystal structure of the 5' exonuclease of bacteriophage T5 showed a distinct arch formed by two helices, one positively charged and one containing hydrophobic residues (Ceska *et al.*, Nature 382:90 [1996]). Interestingly, three residues that are conserved between T5 and Taq, Lys 83, Arg 86 and Tyr 82 are all in the arch. These correspond to amino acids Lys 83, Arg 86, and Tyr 82 in Taq DNA polymerase. The crystal structure for Taq (5' nuclease) has also been determined (Kim *et al.*, Nature 376:612 [1995]).

The crystal structure from the flap endonuclease-1 from *Methanococcus janneschii* also shows such a loop motif (Hwang *et al.*, Nat. Struct. Biol., 5:707 [1998]). The backbone crystal structure of Mja FEN-1 molecules may be superimposed on T5 exonuclease, Taq 5'-exonuclease and T4 RNase H. An interesting feature common to all of these is the long loop. The loop of FEN-1 consists of a number of positively charged and aromatic residues and forms a hole with dimensions large enough to accommodate a single-stranded DNA molecule. The corresponding region in T5 exonuclease consists of three helices forming a helical arch. The size of the hole formed by the helical arch in T5 exonuclease is less than half that formed by the L1 loop in Mj FEN-1. In T4 RNase H or Taq 5' exonuclease, this region is disordered. Some regions of the arch bind metals, while other regions of the arch contact nucleic acid substrate. Alignment of the amino-acid sequences of six 5' nuclease domains from DNA polymerases in the pol I family show six highly conserved sequence motifs containing ten conserved acidic residues (Kim *et al.*, Nature 376 [1995]).

The effects of alterations in the arch region were examined. In Taq polymerase the arch region is formed by amino acids 80-95 and 96-109. Site directed mutagenesis was performed on the arch region. Alignment of amino acid sequences of the FEN and polymerase 5' nucleases suggested the design of 3 amino acid substitution mutations, P88E, P90E and G80E. These substitutions were made on the Taq4M polymerase mutant as a parent enzyme. Results indicated that although the background activity on the HP and X substrates shown in Figure 22 are tremendously suppressed in all mutants, the desirable 5' nuclease activity on proper substrates (IdT and IrT, Figure 24) is also reduced. Despite the sequence homology between Taq and Tth polymerases, they have very different activity on HP and X substrates. The alignment of the Taq and Tth polymerase arch regions also demonstrates regions of extensive sequence homology as well as minor differences. These differences led to the design of mutations L109F and A110T using Taq4M to generate Taq4M L109F/A110T, and the mutant Taq 4M A502K/G504K/E507K/T514S to generate Taq 4M L109F/A110T/A502K/G504K/E507K/T514S mutant. These two mutations have drastically converted Taq4M enzyme to become more like Tth enzyme in terms of the background substrate specificity while the 5' nuclease activities on both DNA and RNA substrates are almost unchanged.

IV) Focused random mutagenesis

As described above, physical studies and molecular modeling may be used alone or in combination to identify regions of the enzymes in which changes of amino acids are likely to cause observable differences in at least one aspect of cleavage function. In the section above, use of this information was described to select and change specific amino acids or combinations of amino acids. Another method of generating an enzyme with altered function is to introduce mutations randomly. Such mutations can be introduced by a number of methods known in the art, including but not limited to, PCR amplification under conditions that favor nucleotide misincorporation (REF), amplification using primers having regions of degeneracy (*i.e.*, base positions in which different individual, but otherwise similar oligonucleotides in a reaction may have different bases), and chemical synthesis. Many methods of random mutagenesis are known in the art (Del Rio *et al.*, Biotechniques 17:1132

[1994]), and may be incorporated into the production of the enzymes of the present invention. The discussions of any particular means of mutagenesis contained herein are presented solely by way of example and not intended as a limitation. When random mutagenesis is performed such that only a particular region of an entire protein is varied, it can be described as "focused random mutagenesis." As described in the Experimental Examples, a focused random mutagenesis approach was applied to vary the HhH and the thumb domains some of the enzyme variants previously created. These domains were chosen to provide examples of this approach, and it is not intended that the random mutagenesis approach be limited to any particular domain, or to a single domain. It may be applied to any domain, or to any entire protein. Proteins thus modified were tested for cleavage activity in the screening reactions described in Example 1, using the test substrates diagrammed in Figures 22 and 24, with the results described in Tables 5 and 6.

Random mutagenesis was performed on the HhH region with the parent TaqSS or TthDN H785A mutants. None of the 8 mutants generated showed an improvement in activity compared to the parent enzyme (Table 5). In fact, mutations of the region between residues 198-205 have about 2-5 fold lower activity on both DNA and RNA substrates, suggesting that this region is essential for substrate recognition. Mutagenesis in the thumb region resulted in new mutations that improved 5' nuclease activity by 20-100% on a DNA target and about 10% on an RNA target (Table 6).

Numerous amino acids in each of the distinct subdomains play roles in substrate contact. Mutagenesis of these may alter substrate specificity by altering substrate binding. Moreover, mutations introduced in amino acids that do not directly contact the substrate may also alter substrate specificity through longer range or general conformation altering effects. These mutations may be introduced by any of several methods known in the art, including, but not limited to random mutagenesis, site directed mutagenesis, and generation of chimeric proteins.

As noted above, numerous methods of random mutagenesis are known in the art. The methods applied in the focused random mutagenesis described herein may be applied to whole genes. It is also contemplated that additional useful chimerical constructs may be created through the use of molecular breeding (*See e.g.*, U.S. Pat. No. 5,837,458 and PCT

Publications WO 00/18906, WO 99/65927, WO 98/31837, and WO 98/27230, herein incorporated by reference in their entireties). Regardless of the mutagenesis method chosen, the rapid screening method described herein provides a fast and effective means of identifying beneficial changes within a large collection of recombinant molecules. This makes the random mutagenesis procedure a manageable and practical tool for creating a large collection of altered 5' nucleases having beneficial improvements. The cloning and mutagenesis strategies employed for the enzymes used as examples are applicable to other thermostable and non-thermostable Type A polymerases, since DNA sequence similarity among these enzymes is very high. Those skilled in the art would understand that differences in sequence would necessitate differences in cloning strategies, for example, the use of different restriction endonucleases may be required to generate chimeras. Selection of existing alternative sites, or introduction via mutagenesis of alternative sites are well established processes and are known to one skilled in the art.

Enzyme expression and purification can be accomplished by a variety of molecular biology methods. The examples described below teach one such method, though it is to be understood that the present invention is not to be limited by the method of cloning, protein expression, or purification. The present invention contemplates that the nucleic acid construct be capable of expression in a suitable host. Numerous methods are available for attaching various promoters and 3' sequences to a gene structure to achieve efficient expression.

EXPERIMENTAL

The following examples serve to illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

In the disclosure which follows, the following abbreviations apply: Afu (*Archaeoglobus fulgidus*); Mth (*Methanobacterium thermoautotrophicum*); Mja (*Methanococcus jannaschii*); Pfu (*Pyrococcus furiosus*); Pwo (*Pyrococcus woesei*); Taq (*Thermus aquaticus*); TaqPol (*T. aquaticus* DNA polymerase I); StfPol (the Stoffel fragment of TaqPol); Ec1Pol (*E. coli* DNA polymerase I); Tth (*Thermus thermophilus*); TthPol (*T. thermophilus* DNA polymerase I); Tsc (*Thermus scotoductus*); TscPol (*Thermus scotoductus*

DNA polymerase I); Tfi (*Thermus filiformis*); TfiPol (*Thermus filiformis* DNA polymerase I);
Ex. (Example); Fig. (Figure); °C (degrees Centigrade); g (gravitational field); vol (volume);
w/v (weight to volume); v/v (volume to volume); BSA (bovine serum albumin); CTAB
(cetyltrimethylammonium bromide); HPLC (high pressure liquid chromatography); DNA
(deoxyribonucleic acid); p (plasmid); µl (microliters); ml (milliliters); µg (micrograms);
pmoles (picomoles); mg (milligrams); M (molar); mM (milliMolar); µM (microMolar); nm
(nanometers); kdal (kilodaltons); OD (optical density); EDTA (ethylene diamine tetra-acetic
acid); FITC (fluorescein isothiocyanate); SDS (sodium dodecyl sulfate); NaPO₄ (sodium
phosphate); Tris (tris(hydroxymethyl)-aminomethane); PMSF (phenylmethylsulfonylfluoride);
TAE (Tris-acetate-EDTA, *i.e.*, 40mM Tris-Acetate, 1 mM EDTA); TBE (Tris-Borate-EDTA,
i.e., Tris buffer titrated with boric acid rather than HCl and containing EDTA); PBS
(phosphate buffered saline); PPBS (phosphate buffered saline containing 1 mM PMSF);
PAGE (polyacrylamide gel electrophoresis); Tween (polyoxyethylene-sorbitan); ATCC
(American Type Culture Collection, Rockville, MD); DSMZ (Deutsche Sammlung von
Mikroorganismen und Zellculturen, Braunschweig, Germany); Sigma (Sigma Chemical
Company, St. Louis, MO); Dynal (Dynal A.S., Oslo, Norway); Gull (Gull Laboratories, Salt
Lake City, UT); Epicentre (Epicentre Technologies, Madison, WI); MJ Research (MJ
Research, Watertown, MA); National Biosciences (Plymouth, MN); New England Biolabs
(Beverly, MA); Novagen (Novagen, Inc., Madison, WI); Perkin Elmer (Norwalk, CT);
Promega Corp. (Madison, WI); Stratagene (Stratagene Cloning Systems, La Jolla, CA);
Clontech (Clontech, Palo Alto, CA); Pharmacia (Pharmacia, Piscataway, NJ); Milton Roy
(Milton Roy, Rochester, NY); Amersham (Amersham International, Chicago, IL); USB (U.S.
Biochemical, Cleveland, OH); and Qiagen (Valencia, CA).

EXAMPLE 1

Rapid screening of colonies for 5' nuclease activity

The native 5' nucleases and the enzymes of the present invention can be tested directly
for a variety of functions. These include, but are not limited to, 5' nuclease activity on RNA
or DNA targets and background specificity using alternative substrates representing structures

that may be present in a target detection reaction. Examples of nucleic acid molecules having suitable test structures are shown schematically in Figures 18A-D and Figures 21-24. The screening techniques described below were developed to quickly and efficiently characterize 5' nucleases and to determine whether the new 5' nucleases have any improved or desired activities. Enzymes that show improved cycling rates on RNA or DNA targets, or that result in reduced target-independent cleavage merit more thorough investigation. In general, the modified proteins developed by random mutagenesis were tested by rapid colony screen on the substrates shown in Figures 18A and 18B. A rapid protein extraction was then done, and a test of activity on alternative structures, (*e.g.*, as shown in Figures 18C-D) was performed using the protein extract. Either the initial screen, or further screening and characterization of enzymes for improved activity may be performed using other cleavage complexes, such as those diagrammed in Figures 21-24. It is not intended that the scope of the invention be limited by the particular sequences used to form such test cleavage structures. One skilled in the art would understand how to design and create comparable nucleic acids to form analogous structures for rapid screening.

This order of testing may be chosen to reduce the number of tests overall, to save time and reagents. The order of testing for enzyme function is not intended to be a limitation on the present invention. Those mutants that showed reasonable cycling rates with the RNA or DNA targets may then be cultured overnight, and a rapid protein extraction done. Alternatively, any subset or all of the cleavage tests may be done at the same time.

For convenience, each type of rapid screen may be done on a separate microtiter plate. For example, one plate may be set up to test RNA INVADER activity, one plate set up to test for DNA INVADER activity. As many as 90 different colonies may be screened on one plate. The colonies screened can be from a variety of sources, such as clones of unaltered (native) 5' nucleases, from one mutagenesis reaction (*e.g.*, many colonies from a single plate) or from a variety of reactions (colonies selected from multiple plates).

Ideally, positive and negative controls should be run on the same plate as the mutants, using the same preparation of reagents. One example of a good positive control is a colony containing the unmodified enzyme, or a previously modified enzyme whose activity is to be compared to new mutants. For example, if a mutagenesis reaction is performed on the Taq

DN RX HT construct (described below), the unmodified Taq DN RX HT construct would be chosen as the standard for comparing the effects of mutagenesis on enzymatic activity. Additional control enzymes may also be incorporated into the rapid screening test. For example, Tth DN RX HT (described below; unless otherwise specified, the TaqPol and TthPol enzymes of the following discussion refer to the DN RX HT derivative) may also be included as a standard for enzymatic activity along with the Taq DN RX HT. This would allow a comparison of any altered enzymes to two known enzymes having different activities. A negative control should also be run to determine the background reaction levels (*i.e.*, cleavage or probe degradation due to sources other than the nucleases being compared). A good negative control colony would be one containing only the vector used in the cloning and mutagenesis, for example, colonies containing only the pTrc99A vector.

Two factors that may influence the number of colonies chosen from a specific mutagenesis reaction for the initial rapid screen are 1) total number of colonies obtained from the mutagenesis reaction, and 2) whether the mutagenesis reaction was site-specific or randomly distributed across a whole gene or a region of a gene. For example, if only 5-10 colonies are present on the plate, all colonies can easily be tested. If hundreds of colonies are present, a subset of these may be analyzed. Generally 10-20 colonies are tested from a site-specific mutagenesis reaction, while 80 to 100 or more colonies are routinely tested from a single random mutagenesis reaction.

Where indicated, the altered 5' nucleases described in these experimental examples were tested as detailed below.

A. Rapid screen: INVADER activity on RNA target (Figure 18A)

A 2X substrate mix was prepared, comprising 20 mM MOPS, pH 7.5, 10 mM MgSO₄, 200 mM KCl, 2 μM FRET-probe oligo SEQ ID NO:21 (5'-Fl-CGCT-cy3-TCTCGCTCGC-3'), 1 μM INVADER oligo SEQ ID NO:22 (5'-ACGGAACGAGCGTCTTTG-3'), and 4 nM RNA target SEQ ID NO:23 (5'-GCG AGC GAGA CAG CGA AAG ACG CUC GUU CCG U-3'). Five μl of the 2X substrate mix were dispensed into each sample well of a 96 well microtiter plate (Low Profile MULTIPLATE 96, M.J. Research, Inc.).

Cell suspensions were prepared by picking single colonies (mutants, positive control, and negative control colonies) and suspending each one in 20µl of water. This can be done conveniently in a 96 well microtiter plate format, using one well per colony.

Five µl of the cell suspension was added to the appropriate test well such that the final reaction conditions were 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl, 1 µM FRET-probe oligo, 0.5 µM INVADER oligo, and 2 nM RNA target. The wells were covered with 10 µl of Clear Chill-out 14 (M.J. Research, Inc.) liquid wax, and the samples were heated at 85°C for 3 minutes, then incubated at 59°C for 1 hour. After the incubation, the plates were read on a Cytofluor fluorescence plate reader using the following parameters: excitation 485/20, emission 530/30.

B. Rapid screen: INVADER activity on DNA target (Figure 18B)

A 2X substrate mix was prepared, comprising 20 mM MOPS, pH 7.5, 10 mM MgSO₄, 200 mM KCl, 2 µM FRET-probe oligo SEQ ID NO:21 (5'-F1-CGCT-Cy3-TCTCGCTCGC-3'), 1 µM INVADER oligo SEQ ID NO:22 (5'-ACGGAACGAGCGTCTTTG-3'), 1 nM DNA target SEQ ID NO:24 (5'-GCG AGC GAGA CAG CGA AAG ACG CTC GTT CCG T-3'). Five µl of the 2X substrate mix was dispensed into each sample well of a 96 well microtiter plate (MJ Low Profile).

Cell suspensions were prepared by picking single colonies (mutants, positive control and negative control colonies) and suspending them in 20 µl of water, generally in a 96 well microtiter plate format.

5 µl of the cell suspension were added to the appropriate test well such that the final reaction conditions were 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl, 1 µM FRET-probe oligo, 0.5 µM INVADER oligo, and 0.5 nM DNA target. Wells were covered with 10µ l of Clear Chill-out 14 (M.J. Research, Inc.) liquid wax, and the reactions were heated at 85°C for 3 minutes, then incubated at 59°C for 1 hour. After the hour incubation, the plate were read on a Cytofluor fluorescence plate reader using the following parameters: excitation 485/20, emission 530/30, gain 40, reads per well 10.

C. Rapid protein extraction (crude cell lysate)

Those mutants that gave a positive or an unexpected result in either the RNA or DNA INVADER assay were further analyzed, specifically for background activity on the X-structure or the hairpin substrate (Figure 18C and D, respectively). A rapid colony screen format can be employed, as described above. By simply changing the substrate, tests for background or aberrant enzymatic activity can be done. Another approach would be to do a rapid protein extraction from a small overnight culture of positive clones, and then test this crude cell lysate for additional protein function. One possible rapid protein extraction procedure is detailed below. Two to five ml of LB (containing the appropriate antibiotic for plasmid selection; *See e.g.*, Maniatis, books 1,2 and 3) were inoculated with the remaining volume of the 20 μ l water-cell suspension and incubated at 37°C overnight. About 1.4 ml of the culture were transferred to a 1.5 ml microcentrifuge tube, and microcentrifuged at top speed (*e.g.*, 14,000 rpm in an Eppendorf 5417 table top microcentrifuge), at room temperature for 3-5 minutes to pellet the cells. The supernatant was removed, and the cell pellet was suspended in 100 μ l of TES buffer pH 7.5 (Sigma). Lysozyme (Promega) was added to a final concentration of 0.5 μ g/ μ l and samples were incubated at room temperature for 30 minutes. Samples were then heated at 70°C for 10 minutes to inactivate the lysozyme, and the cell debris was pelleted by microcentrifugation at top speed for 5 minutes. The supernatant was removed and this crude cell lysate was used in the following enzymatic activity assays.

D. Rapid screen: background specificity X structure substrate (Figure 18C)

Reactions were performed under conditions as detailed above. One μ l of crude cell lysate was added to 9 μ l of reaction components for a final volume of 10 μ l and final concentrations of 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl, 1 μ M FRET-probe oligo (SEQ ID NO:21), 0.5 μ M X-structure INVADER oligo SEQ ID NO:25 (5'-ACGGAACGAGCGTCTTTCATCTGTCAATC-3'), and 0.5 nM DNA target (SEQ ID NO:24). Wells were covered with 10 μ l of Clear Chill-out 14 (M.J. Research, Inc.) liquid wax, and the reactions were heated at 85°C for 3 minutes, then incubated at 59°C for 1 hour. After the incubation, the plates were read on a Cytofluor fluorescence plate reader using the

following parameters: excitation 485/20, emission 530/30, gain 40, reads per well 10.

E. Rapid screen: background specificity hairpin substrate (Figure 18D)

Reactions were performed under conditions as detailed above. One μl of crude cell
lysate was added to 9 μl of reaction components for a final volume of 10 μl and final
concentrations of 10 mM MOPS, pH 7.5, 5 mM MgSO_4 , 100 mM KCl, 1 μM FRET-probe
oligonucleotide (SEQ ID NO:21), and 0.5 nM DNA target (SEQ ID NO:24). Wells were
covered with 10 μl of Clear Chill-out 14 (M.J. Research, Inc.) liquid wax, and the reactions
were heated at 85°C for 3 minutes, then incubated at 59°C for 1 hour. After the hour
incubation, the plate were read on a Cytofluor plate reader using the following parameters:
excitation 485/20, emission 530/30, gain 40, reads per well 10.

F. Activity assays with IrT1 and IdT targets (Figures 24)

The 5' nuclease activities assays were carried out in 10 μl of a reaction containing 10
mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 $\mu\text{g/ml}$ tRNA, 100 mM KCl
and 5 mM MgSO_4 . The probe concentration (SEQ ID NO: 26) was 2 mM. The substrates
(IrT1 (SEQ ID NO: 35) or IdT (SEQ ID NO: 36) at 10 or 1 nM final concentration
respectively) and approximately 20 ng of an enzyme, prepared as in Example 3, were mixed
with the above reaction buffer and overlaid with Chill-out (MJ Research) liquid wax.
Reactions were brought up to reaction temperature 57 °C, started by addition of MgSO_4 , and
incubated for 10 min. Reactions were then stopped by the addition of 10 μl of 95%
formamide containing 10 mM EDTA and 0.02% methyl violet (Sigma). Samples were heated
to 90°C for 1 minute immediately before electrophoresis through a 20% denaturing acrylamide
gel (19:1 cross-linked), with 7 M urea, and in a buffer of 45 mM Tris-borate, pH 8.3, 1.4 mM
EDTA. Unless otherwise indicated, 1 μl of each stopped reaction was loaded per lane. Gels
were then scanned on an FMBIO-100 fluorescent gel scanner (Hitachi) using a 505 nm filter.
The fraction of cleaved product was determined from intensities of bands corresponding to
uncut and cut substrate with FMBIO Analysis software (version 6.0, Hitachi). The fraction of
cleavage product did not exceed 20% to ensure that measurements approximated initial
cleavage rates. The turnover rate was defined as the number of cleaved signal probes

generated per target molecule per minute under these reaction conditions (1/min).

G. Activity assays with X structure (X) and hairpin (HP) targets (Figures 22)

The 5' nuclease activity assays were carried out in 10 μ l of a reaction containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 μ g/ml tRNA, 100 mM KCl and 5 mM MgSO₄. Each oligo for formation of either the hairpin structure assembly (22A, SEQ ID NOS: 29 and 30) assembly or the X structure assembly (22B, SEQ ID NOS: 29-31) was added to a final concentration of 1 μ M, and approximately 20 ng of test enzyme prepared as described in Example 3, were mixed with the above reaction buffer and overlaid with Chill-out (MJ Research) liquid wax. Reactions were brought up to reaction temperature 60 °C, started by addition of MgSO₄, and incubated for 10 min. Reactions were then stopped by the addition of 10 μ l of 95% formamide containing 10 mM EDTA and 0.02% methyl violet (Sigma). Samples were heated to 90°C for 1 minute immediately before electrophoresis through a 20% denaturing acrylamide gel (19:1 cross-linked), with 7 M urea, and in a buffer of 45 mM Tris-borate, pH 8.3, 1.4 mM EDTA. Unless otherwise indicated, 1 μ l of each stopped reaction was loaded per lane. Gels were then scanned on an FMBIO-100 fluorescent gel scanner (Hitachi) using a 505 nm filter. The fraction of cleaved product was determined from intensities of bands corresponding to uncut and cut substrate with FMBIO Analysis software (version 6.0, Hitachi). The fraction of cleavage product did not exceed 20% to ensure that measurements approximated initial cleavage rates. The turnover rate was defined as the number of cleaved signal probes generated per target molecule per minute under these reaction conditions (1/min).

H. Activity assays with human IL-6 target (Figure 10)

The 5' nuclease activities assays were carried out in 10 μ l reactions containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 μ g/ml tRNA, 100 mM KCl and 5 mM MgSO₄. Reactions comprising the DNA IL-6 substrate contained 0.05 nM IL-6 DNA target (SEQ ID NO: 18) and 1 μ M of each probe (SEQ ID NO: 16) and INVADER (SEQ ID NO: 15) oligonucleotides, and were carried out at 60°C for 30 min. Reactions comprising the IL-6 RNA target (SEQ ID NO: 17) were performed under the same

conditions, except that the IL-6 RNA target concentration was 1 nM and the reactions were performed at 57°C for 60 min. Each reaction contained approximately 20 ng of test enzyme, prepared as described in Example 3.

I. Activity assays with synthetic r25mer target (Figure 23)

Reactions comprising the synthetic r25mer target (SEQ ID NO: 34) were carried out under the same reaction conditions (10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 µg/ml tRNA, 100 mM KCl and 5 mM MgSO₄) and 1 µM of each probe (SEQ ID NO: 32) and INVADER (SEQ ID NO: 33) oligonucleotides, except that the r25mer target concentration was 5 nM and the reactions were performed at 58°C for 60 min. Approximately 20 ng of each test enzyme was added to the reactions. Enzymes were prepared as described in Example 3.

Any of the tests described above can be modified to derive the optimal conditions for enzymatic activity. For example, enzyme titrations can be done to determine the optimal enzyme concentration for maximum cleavage activity, and lowest background signal. By way of example, but not by way of limitation, many of the mutant enzymes were tested at 10, 20 and 40 ng amounts. Similarly, a temperature titration can also be incorporated into the tests. Since modifying the structure of a protein can alter its temperature requirements, a range of temperatures can be tested to identify the condition best suited for the mutant in question.

Examples of the results from such screens (using approximately 20 ng of the mutant enzyme) are shown in Tables 2-7, and Figures 12, 14, 15, 19, and 25.

EXAMPLE 2

Cloning and Expression of 5' nucleases of DNA polymerases and mutant polymerases

A. DNA polymerases of *Thermus aquaticus* and *Thermus thermophilus*

1. Cloning of TaqPol and TthPol

Type A DNA polymerases from eubacteria of the genus *Thermus* share extensive

protein sequence identity (90% in the polymerization domain, using the Lipman-Pearson method in the DNA analysis software from DNASTar, WI) and behave similarly in both polymerization and nuclease assays. Therefore, the genes for the DNA polymerase of *Thermus aquaticus* (TaqPol), *Thermus thermophilus* (TthPol) and *Thermus scotoductus* were used as representatives of this class. Polymerase genes from other eubacterial organisms, including, but not limited to, *Escherichia coli*, *Streptococcus pneumoniae*, *Mycobacterium smegmatis*, *Thermus thermophilus*, *Thermus sp.*, *Thermotoga maritima*, *Thermosiphon africanus*, and *Bacillus stearothermophilus* are equally suitable.

a. Initial TaqPol Isolation: mutant TaqA/G

The *Taq* DNA polymerase gene was amplified by polymerase chain reaction from genomic DNA from *Thermus aquaticus*, strain YT-1 (Lawyer *et al.*, *supra*), using as primers the oligonucleotides described in SEQ ID NOS:37 and 38. The resulting fragment of DNA has a recognition sequence for the restriction endonuclease *Eco*RI at the 5' end of the coding sequence and a *Bgl*II sequence at the 3' end of the coding strand. Cleavage with *Bgl*II leaves a 5' overhang or "sticky end" that is compatible with the end generated by *Bam*HI. The PCR-amplified DNA was digested with *Eco*RI and *Bam*HI. The 2512 bp fragment containing the coding region for the polymerase gene was gel purified and then ligated into a plasmid that contains an inducible promoter.

In one embodiment of the invention, the pTTQ18 vector, which contains the hybrid *trp-lac (tac)* promoter, was used (M.J.R. Stark, *Gene* 5:255 [1987]). The *tac* promoter is under the control of the *E. coli lac* repressor protein. Repression allows the synthesis of the gene product to be suppressed until the desired level of bacterial growth has been achieved, at which point repression is removed by addition of a specific inducer, isopropyl-b-D-thiogalactopyranoside (IPTG). Such a system allows the controlled expression of foreign proteins that may slow or prevent growth of transformants.

Particularly strong bacterial promoters, such as the synthetic *Ptac*, may not be adequately suppressed when present on a multiple copy plasmid. If a highly toxic protein is placed under control of such a promoter, the small amount of expression leaking through,

even in the absence of an inducer, can be harmful to the bacteria. In another embodiment of the invention, another option for repressing synthesis of a cloned gene product is contemplated. A non-bacterial promoter from bacteriophage T7, found in the plasmid vector series pET-3, was used to express the cloned mutant *Taq* polymerase genes (Studier and Moffatt, J. Mol. Biol., 189:113 [1986]). This promoter initiates transcription only by T7 RNA polymerase. In a suitable strain, such as BL21(DE3)pLYS, the gene for the phage T7 RNA polymerase is carried on the bacterial genome under control of the *lac* operator. This arrangement has the advantage that expression of the multiple copy gene (on the plasmid) is completely dependent on the expression of T7 RNA polymerase, which is easily suppressed because it is present in a single copy.

These are just two examples of vectors having suitable inducible promoters. Others are well known to those skilled in the art, and it is not intended that the improved nucleases of the present invention be limited by the choice of expression system.

For ligation into the pTTQ18 vector, the PCR product DNA containing the *Taq* polymerase coding region (termed mut*Taq* for reasons discussed below, SEQ ID NO:39) was digested with *Eco*RI and *Bgl*II and this fragment was ligated under standard "sticky end" conditions (Sambrook *et al. Molecular Cloning*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 1.63-1.69 [1989]) into the *Eco*RI and *Bam*HI sites of the plasmid vector pTTQ18. Expression of this construct yields a translational fusion product in which the first two residues of the native protein (Met-Arg) are replaced by three from the vector (Met-Asn-Ser), but the remainder of the PCR product's protein sequence is not changed (SEQ ID NO:40). The construct was transformed into the JM109 strain of *E. coli*, and the transformants were plated under incompletely repressing conditions that do not permit growth of bacteria expressing the native protein. These plating conditions allow the isolation of genes containing pre-existing mutations, such as those that result from the infidelity of *Taq* polymerase during the amplification process.

Using this amplification/selection protocol, a clone was isolated containing a mutated *Taq* polymerase gene (mut*Taq*). The mutant was first detected by its phenotype, in which temperature-stable 5' nuclease activity in a crude cell extract was normal, but polymerization

activity was almost absent (approximately less than 1% of wild type *Taq* polymerase activity). Polymerase activity was determined by primer extension reactions. The reactions were carried out in 10 µl of buffer containing 10 mM MOPS, pH 7.5, 5 mM MgSO₄, 100 mM KCl. In each reaction, 40 ng of enzyme were used to extend 10 µM (dT)₂₅₋₃₀ primer in the presence of either 10 µM poly (A)₂₈₆ or 1 µM poly (dA)₂₇₃ template, 45 µM dTTP and 5 µM Fl-dUTP at 60°C for 30 minutes. Reactions were stopped with 10 µl of stop solution (95% formamide, 10 mM EDTA, 0.02% methyl violet dye). Samples (3 µl) were fractionated on a 15% denaturing acrylamide gel (19:1 crossed-linked) and the fraction of incorporated Fl-dUTP was quantitated using an FMBIO-100 fluorescence gel scanner (Hitachi) equipped with a 505 nm emission filter.

DNA sequence analysis of the recombinant gene showed that it had changes in the polymerase domain resulting in two amino acid substitutions: an A to G change at nucleotide position 1394, which causes a Glu to Gly change at amino acid position 465 (numbered according to the natural nucleic and amino acid sequences, SEQ ID NOS:1 and 4), and another A to G change at nucleotide position 2260, which causes a Gln to Arg change at amino acid position 754. Because the Gln to Gly mutation is at a nonconserved position and because the Glu to Arg mutation alters an amino acid that is conserved in virtually all of the known Type A polymerases, the latter mutation is most likely the one responsible for curtailing the synthesis activity of this protein. The nucleotide sequence for the construct is given in SEQ ID NO:39. The enzyme encoded by this sequence is referred to as *Taq* A/G.

b. Initial TthPol Isolation

The DNA polymerase enzyme from the bacterial species *Thermus thermophilus* (Tth) was produced by cloning the gene for this protein into an expression vector and overproducing it in *E. coli* cells. Genomic DNA was prepared from 1 vial of dried *Thermus thermophilus* strain HB-8 from ATCC (ATCC #27634). The DNA polymerase gene was amplified by PCR using the following primers:

5'-CACGAATTCCGAGGCGATGCTTCCGCTC-3' (SEQ ID NO:41) and

5'-TCGACGTCGACTAACCCTTGCGGAAAGCC-3' (SEQ ID NO:42). The resulting PCR

product was digested with *EcoRI* and *SalI* restriction endonucleases and inserted into *EcoRI/Sal I* digested plasmid vector pTrc99G (described in Example 2C1) to create the plasmid pTrcTth-1. This *Tth* polymerase construct is missing a single nucleotide that was inadvertently omitted from the 5' oligonucleotide, resulting in the polymerase gene being out of frame. This mistake was corrected by site specific mutagenesis of pTrcTth-1 as described in Examples 4 and 5 using the following oligonucleotide:
5'-GCATCGCCTCGGAATTCATGGTC-3' (SEQ ID NO:43), to create the plasmid pTrcTth-2. The protein and the nucleic acid sequence encoding the protein are referred to as TthPol, and are listed as SEQ ID NOS:6 and 3 respectively.

c. Large Scale preparation of recombinant proteins

The recombinant proteins were purified by the following technique which is derived from a *Taq* DNA polymerase preparation protocol (Engelke *et al.*, Anal. Biochem., 191:396 [1990]) as follows. *E. coli* cells (strain JM109) containing either pTrc99A TaqPol, pTrc99GTthPol were inoculated into 3 ml of LB containing 100 mg/ml ampicillin and grown for 16 hrs at 37°C. The entire overnight culture was inoculated into 200 ml or 350 ml of LB containing 100 mg/ml ampicillin and grown at 37°C with vigorous shaking to an A_{600} of 0.8. IPTG (1 M stock solution) was added to a final concentration of 1 mM and growth was continued for 16 hrs at 37°C.

The induced cells were pelleted and the cell pellet was weighed. An equal volume of 2X DG buffer (100 mM Tris-HCl, pH 7.6, 0.1 mM EDTA) was added and the pellet was suspended by agitation. Fifty mg/ml lysozyme (Sigma) were added to 1 mg/ml final concentration and the cells incubated at room temperature for 15 min. Deoxycholic acid (10% solution) was added dropwise to a final concentration of 0.2 % while vortexing. One volume of H₂O and 1 volume of 2X DG buffer were added, and the resulting mixture was sonicated for 2 minutes on ice to reduce the viscosity of the mixture. After sonication, 3 M (NH₄)₂SO₄ was added to a final concentration of 0.2 M, and the lysate was centrifuged at 14000 x g for 20 min at 4°C. The supernatant was removed and incubated at 70°C for 60 min at which time 10% polyethylimine (PEI) was added to 0.25%. After incubation on ice

for 30 min., the mixture was centrifuged at 14,000 x g for 20 min at 4°C. At this point, the supernatant was removed and the protein precipitated by the addition of (NH₄)₂SO₄ as follows.

Two volumes of 3 M (NH₄)₂SO₄ were added to precipitate the protein. The mixture was incubated overnight at room temperature for 16 hrs centrifuged at 14,000 x g for 20 min at 4°C. The protein pellet was suspended in 0.5 ml of Q buffer (50 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 0.1% Tween 20). For the *Mja* FEN-1 preparation, solid (NH₄)₂SO₄ was added to a final concentration of 3 M (~75% saturated), the mixture was incubated on ice for 30 min, and the protein was spun down and suspended as described above.

The suspended protein preparations were quantitated by determination of the A₂₇₉ dialyzed and stored in 50% glycerol, 20 mM Tris HCl, pH8.0, 50 mM KCl, 0.5% Tween 20, 0.5% Nonidet P-40, with 100 µg/ml BSA.

B. DNA polymerases of *Thermus filiformis* and *Thermus scotoductus*

1. Cloning of *Thermus filiformis* and *Thermus scotoductus*

One vial of lyophilized *Thermus filiformis* (Tfi) obtained from DSMZ (Deutsche Sammlung von Mikroorganismen und Zellculturen, Braunschweig, Germany, strain #4687) was rehydrated in 1 ml of Castenholz medium (DSMZ medium 86) and inoculated into 500 ml of Castenholz medium preheated to 50°C. The culture was incubated at 70°C with vigorous shaking for 48 hours. After growth, the cells were harvested by centrifugation at 8000 x g for 10 minutes, the cell pellet was suspended in 10 ml of TE (10 mM TrisHCL, pH 8.0, 1 mM EDTA), and the cells were frozen at -20°C in 1 ml aliquots. A 1 ml aliquot was thawed, lysozyme was added to 1 mg/ml, and the cells were incubated at 23°C for 30 minutes. A solution of 20% SDS (sodium dodecyl sulfate) was added to a final concentration of 0.5% followed by extraction with buffered phenol. The aqueous phase was further extracted with 1:1 phenol:chloroform, and extracted a final time with chloroform. One-tenth volume of 3 M sodium acetate, pH 5.0 and 2.5 volumes of ethanol were added to the aqueous phase and mixed. The DNA was pelleted by centrifugation at 20,000 x g for 5 minutes. The DNA pellet was washed with 70% ethanol, air dried and resuspended in 200 µl of TE and used directly for amplification. *Thermus scotoductus* (Tsc, ATCC # 51532) was grown and

genomic DNA was prepared as described above for *Thermus filiformis*.

The DNA polymerase I gene from *Tfi* (GenBank accession #AF030320) could not be amplified as a single fragment. Therefore, it was cloned in 2 separate fragments into the expression vector pTrc99a. The 2 fragments overlap and share a *Not I* site which was created by introducing a silent mutation at position 1308 of the *Tfi* DNA polymerase open reading frame (ORF) in the PCR oligonucleotides. The 3' half of the gene was amplified using the Advantage cDNA PCR kit (Clontech) with the following oligonucleotides; 5'-ATAGCCATGGTGGAGCGGCCGCTCTCCCGG (SEQ ID NO:44) and 5'-AAGCGTCGACTCAATCCTGCTTCGCCTCCAGCC (SEQ ID NO:45). The PCR product from this reaction was approximately 1200 base pairs in length. It was cut with the restriction enzymes *Not I* and *Sal I*, and the resulting DNA was ligated into pTrc99a cut with *NotI* and *SalI* to create pTrc99a-Tfi3'. The 5' half of the gene was amplified as described above using the following two primers; 5'-AATCGAATTCACCCCACTTTTGTGACCTGGAGG (SEQ ID NO:46) and 5'-CCGGGAGAGCGGCCGCTCCAC (SEQ ID NO:47). The resulting 1300 base pair fragment was cut with restriction enzymes *Eco RI* and *Not I* and ligated into pTrc99a-Tfi3' cut with *NotI* and *EcoRI* to produce pTrc99a-TfiPol, SEQ ID NO:48 (the corresponding amino acid sequence is listed in SEQ ID NO:155).

The DNA polymerase I gene from *Thermus scotoductus* was amplified using the Advantage cDNA PCR kit (Clontech) using the following two primers; 5'-ACTGGAATTCCTGCCCCTCTTTGAGCCCAAG (SEQ ID NO:49) and 5'-AACAGTCGACCTAGGCCTTGCGGAAAGCC (SEQ ID NO:50). The PCR product was cut with restriction enzymes *Eco RI* and *Sal I* and ligated into *Eco RI*, *Sal I* cut pTrc99a to create pTrc99a-TscPol SEQ ID NO:51 (the corresponding amino acid sequence is listed in SEQ ID NO:268).

2. Expression and purification of *Thermus filiformis* and *Thermus scotoductus*

Plasmids were transformed into protease deficient *E. coli* strain BL21 (Novagen) or strain JM109 (Promega Corp., Madison, WI) for protein expression. Flasks containing 200 ml of LB containing 100 µg/ml ampicillin were inoculated with either a single colony from an

LB plate or from a frozen stock of the appropriate strain. After several hours of growth at 37°C with vigorous shaking, cultures was induced by the addition of 200 µl of 1 M isothiopropyl-galatoside (IPTG). Growth at 37°C was continued for 16 hours prior to harvest. Cells were pelleted by centrifugation at 8000 x g for 15 minutes followed by suspension of the cell pellet in 5 ml of TEN (10 mM TrisHCl, pH 8.0, 1 mM EDTA, 100 mM NaCl). 100 µl of 50 mg/ml lysozyme were added and the cells incubated at room temperature for 15 minutes. Deoxycholic acid (10%) was added to a final concentration of 0.2%. After thorough mixing, the cell lysates were sonicated for 2 minutes on ice to reduce the viscosity of the mixture. Cellular debris was pelleted by centrifugation at 4°C for 15 minutes at 20,000 x g. The supernatant was removed and incubated at 70°C for 30 min after which 10% polyethylimine (PEI) was added to 0.25%. After incubation on ice for 30 minutes, the mixture was centrifuged at 20,000 x g for 20 min at 4°C. At this point, the supernatant containing the enzyme was removed, and the protein was precipitated by the addition of 1.2 g of ammonium sulfate and incubation at 4°C for 1 hour. The protein was pelleted by centrifugation at 4°C for 10 minutes at 20,000 x g. The pellet was resuspended in 4 ml of HPLC Buffer A (50 mM TrisHCl, pH 8.0, 1 mM EDTA). The protein was further purified by affinity chromatography using an Econo-Pac heparin cartridge (Bio-Rad) and a Dionex DX 500 HPLC instrument. Briefly, the cartridge was equilibrated with HPLC Buffer A, and the enzyme extract was loaded on the column and eluted with a linear gradient of NaCl (0-2 M) in the same buffer. Pure protein elutes between 0.5 and 1 M NaCl. The enzyme peak was collected and dialyzed in 50% glycerol, 20 mM Tris HCl, pH 8, 50 mM KCl, 0.5% Tween 20, 0.5% Nonidet P40, 100 mg/ml BSA.

C. Generation of polymerase mutants with reduced polymerase activity but unaltered 5' nuclease activity

All mutants generated in section C were expressed and purified as described in Example 2A1C.

1. Modified TaqPol Genes: TaqDN

A polymerization deficient mutant of *Taq* DNA polymerase called TaqDN was constructed. TaqDN nuclease contains an asparagine residue in place of the wild-type aspartic acid residue at position 785 (D785N).

DNA encoding the TaqDN nuclease was constructed from the gene encoding the Taq A/G in two rounds of site-directed mutagenesis. First, the G at position 1397 and the G at position 2264 of the Taq A/G gene (SEQ ID NO:39) were changed to A at each position to recreate a wild-type TaqPol gene. In a second round of mutagenesis, the wild type TaqPol gene was converted to the Taq DN gene by changing the G at position 2356 to A. These manipulations were performed as follows.

DNA encoding the Taq A/G nuclease was recloned from pTTQ18 plasmid into the pTrc99A plasmid (Pharmacia) in a two step procedure. First, the pTrc99A vector was modified by removing the G at position 270 of the pTrc99A map, creating the pTrc99G cloning vector. To this end, pTrc99A plasmid DNA was cut with *Nco*I and the recessive 3' ends were filled-in using the Klenow fragment of *E.coli* polymerase I in the presence of all four dNTPs at 37°C for 15 min. After inactivation of the Klenow fragment by incubation at 65°C for 10 min, the plasmid DNA was cut with *Eco*RI and the ends were again filled-in using the Klenow fragment in the presence of all four dNTPs at 37°C for 15 min. The Klenow fragment was then inactivated by incubation at 65°C for 10 min. The plasmid DNA was ethanol precipitated, recircularized by ligation, and used to transform *E.coli* JM109 cells (Promega). Plasmid DNA was isolated from single colonies, and deletion of the G at position 270 of the pTrc99A map was confirmed by DNA sequencing.

In a second step, DNA encoding the Taq A/G nuclease was removed from the pTTQ18 plasmid using *Eco*RI and *Sa*II and the DNA fragment carrying the Taq A/G nuclease gene was separated on a 1% agarose gel and isolated with GeneClean II Kit (Bio 101, Vista, CA). The purified fragment was ligated into the pTrc99G vector which had been cut with *Eco*RI and *Sa*II. The ligation mixture was used to transform competent *E.coli* JM109 cells (Promega). Plasmid DNA was isolated from single colonies and insertion of the Taq A/G nuclease gene was confirmed by restriction analysis using *Eco*RI and *Sa*II.

Plasmid DNA pTrcAG carrying the Taq A/G nuclease gene cloned into the pTrc99A vector was purified from 200 ml of JM109 overnight culture using QIAGEN Plasmid Maxi kit (QIAGEN, Chatsworth, CA) according to manufacturer's protocol. pTrcAG plasmid DNA was mutagenized using two mutagenic primers, E465 (SEQ ID NO:52) (Integrated DNA Technologies, Iowa) and R754Q (SEQ ID NO:53) (Integrated DNA Technologies), and the selection primer Trans Oligonucleotide AlwNI/SpeI (Clontech, Palo Alto, CA, catalog #6488-1) according to TRANSFORMER Site-Directed Mutagenesis Kit protocol (Clontech, Palo Alto, CA) to produce a restored wild-type TaqPol gene (pTrcWT).

pTrcWT plasmid DNA carrying the wild-type TaqPol gene cloned into the pTrc99A vector was purified from 200 ml of JM109 overnight culture using QIAGEN Plasmid Maxi kit (QIAGEN, Chatsworth, CA) according to manufacturer's protocol. pTrcWT was then mutagenized using the mutagenic primer D785N (SEQ ID NO:54) (Integrated DNA Technologies) and the selection primer Switch Oligonucleotide SpeI/AlwNI (Clontech, Palo Alto, CA, catalog #6373-1) according to TRANSFORMER Site-Directed Mutagenesis Kit protocol (Clontech, Palo Alto, CA) to create a plasmid containing DNA encoding the Taq DN nuclease. The DNA sequence encoding the Taq DN nuclease is provided in SEQ ID NO:55; the amino acid sequence of Taq DN nuclease is provided in SEQ ID NO:56.

2. Modified TthPol Gene: Tth DN

The *Tth* DN construct was created by mutating the TthPol described above. The sequence encoding an aspartic acid at position 787 was changed by site-specific mutagenesis as described above to a sequence encoding asparagine. Mutagenesis of pTrcTth-2 with the following oligonucleotide: 5'-CAGGAGGAGCTCGTTGTGGACCTGGA-3' (SEQ ID NO:57) was performed to create the plasmid pTrcTthDN. The mutant protein and protein coding nucleic acid sequence is termed TthDN SEQ ID NOS:58 and 59 respectively.

3. Taq DN HT and Tth DN HT

Six amino acid histidine tags (his-tags) were added onto the carboxy termini of Taq DN and Tth DN. The site-directed mutagenesis was performed using the TRANSFORMER Site Directed Mutagenesis Kit (Clontech) according to the manufacturer's instructions. The

EXAMPLE 3

RNA-dependent 5' nuclease activity of TthPol can be conferred on TaqPol by transfer of the N-terminal portion of the DNA polymerase domain

A. Preparation and purification of substrate structures having either a DNA or an RNA target strand

The downstream (SEQ ID NO:16) and upstream probes (SEQ ID NO:15) and the IL-6 DNA (SEQ ID NO:18) (figure 10) target strand were synthesized on a PerSeptive Biosystems instrument using standard phosphoramidite chemistry (Glen Research). The synthetic RNA-DNA chimeric IrT target labeled with biotin at the 5'-end (figure 20A) was synthesized utilizing 2'-ACE RNA chemistry (Dharmacon Research). The 2'-protecting groups were removed by acid-catalyzed hydrolysis according to the manufacturer's instructions. The downstream probes labeled with 5'-fluorescein (Fl) or 5'-tetrachloro-fluorescein (TET) at their 5' ends were purified by reverse phase HPLC using a Resource Q column (Amersham-Pharmacia Biotech). The 648-nucleotide IL-6 RNA target (SEQ ID NO:17) (figure 10) was synthesized by T7 RNA polymerase runoff-transcription of the cloned fragment of human IL-6 cDNA (nucleotides 64-691 of the sequence published in May *et al.*, Proc. Natl. Acad. Sci., 83:8957 [1986]) using a Megascript Kit (Ambion). All oligonucleotides were finally purified by separation on a 20% denaturing polyacrylamide gel followed by excision and elution of the major band. Oligonucleotide concentration was determined by measuring absorption at 260 nm. The biotin labeled IrT target was incubated with a 5-fold excess of streptavidin (Promega) in a buffer containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% NP-40 and 10 µg/ml tRNA at room temperature for 10 min.

B. Introduction of restriction sites to make chimeras

The restriction sites used for formation of chimerical proteins, described below, were chosen for convenience. The restriction sites in the following example have been strategically placed to surround regions shown by crystal structure and other analysis to be functional domains (*See*, Figures 6, 7, and 19). Different sites, either naturally occurring or created via directed mutagenesis can be used to make similar constructs with other Type A polymerase

genes from related organisms. It is desirable that the mutations all be silent with respect to protein function. By studying the nucleic acid sequence and the amino acid sequence of the protein, one can introduce changes in the nucleic acid sequence that have no effect on the corresponding amino acid sequence. If the nucleic acid change required affects an amino acid, one can make the alteration such that the new amino acid has the same or similar characteristics of the one replaced. If neither of these options is possible, one can test the mutant enzyme for function to determine if the nucleic acid alteration caused a change in protein activity, specificity or function. It is not intended that the invention be limited by the particular restriction sites selected or introduced for the creation of the improved enzymes of the present invention.

C. Generation of Tth DN RX HT and Taq DN RX HT

Mutagenesis was performed to introduce 3 additional, unique restriction sites into the polymerase domain of both the Taq DN HT and Tth DN HT enzymes. Site specific mutagenesis was performed using the Transformer Site-Directed Mutagenesis Kit from (Clontech) according to manufacturer's instructions. One of two different selection primers, Trans Oligo AlwNI/SpeI or Switch Oligo SpeI/AlwNI (Clontech, Palo Alto CA catalog #6488-1 or catalog #6373-1) was used for all mutagenesis reactions described. The selection oligo used in a given reaction is dependent on the selection restriction site present in the vector. All mutagenic primers were synthesized by standard synthetic chemistry. Resultant colonies were expressed in *E.coli* strain JM109.

The Not I sites (amino acid position 328) were created using the mutagenic primers 5'-gccgccagggcgccgcgtccaccgggcc (SEQ ID NO:66) and 5'-gcctgcagggcgccgcgtgcaccggggca (SEQ ID NO:67) corresponding to the sense strands of the Taq DN HT and the Tth DN HT genes, respectively. The BstI (amino acid position 382) and NdeI (amino acid position 443) sites were introduced into both genes using sense strand mutagenic primes 5'-ctcctggacccttcgaacaccacccc (SEQ ID NO:68) and 5'-gtcctggcccatatggaggccac (SEQ ID NO:69). The mutant plasmids were over-expressed and purified using Qiagen QiaPrep Spin Mini Prep Kit (cat. # 27106). The vectors were tested for the presence of the restriction sites by DNA sequencing and restriction mapping. These

constructs are termed Tth DN RX HT (DNA sequence SEQ ID NO:70; amino acid sequence SEQ ID NO:72) and Taq DN RX HT (DNA sequence SEQ ID NO:71; amino acid sequence SEQ ID NO:73).

D. Chimeras

The chimeric constructs shown in Figure 19 were created by exchanging homologous DNA fragments defined by the restriction endonuclease sites EcoRI (E) and BamHI (B), common for both genes, the cloning vector site SalI (S) and the new sites, NotI (N), BstBI (Bs), NdeI (D) created at the homologous positions of both genes by site directed mutagenesis. In generating these chimeric enzymes, two different pieces of DNA are ligated together to yield the final construct. The larger piece of DNA that contains the plasmid vector as well as part of the Taq or Tth (or parts of both) sequence will be termed the “vector.” The smaller piece of DNA that contains sequences of either the Taq or Tth (or parts of both) polymerase will be termed the “insert.”

All restriction enzymes were from New England Biolabs or Promega and used in reactions with the accompanying buffer, according to the manufacturer’s instructions. Reactions were done in 20 µl volume with about 500 ng of DNA per reaction, at the optimal temperature for the specified enzyme. More than one enzyme was used in a single reaction (double digest) if the enzymes were compatible with respect to reaction buffer conditions and reaction temperature. If the enzymes in question were not compatible with respect to buffer conditions, the enzyme requiring the lowest salt condition was used first. After the completion of that reaction, buffer conditions were changed to be optimal or better suited to the second enzyme, and the second reaction was performed. These are common restriction enzyme digest strategies, well known to those in the art of basic molecular biology (Maniatis, supra).

The digested restriction fragments were gel isolated for optimal ligation efficiency. Two µl of 10X loading dye (50% glycerol, 1X TAE, 0.5% bromophenol blue) were added to the 20 µl reaction. The entire volume was loaded and run on a 1%, 1X TAE agarose gel containing 1 µl of a 1% ethidium bromide solution per 100 ml of agarose gel solution. The digested fragments were visualized under UV light, and the appropriate fragments (as

determined by size) were excised from the gel. These fragments were then purified using the Qiagen Gel Extractio Kit, (cat # 28706) according to the manufacturer's instructions.

Ligations were performed in a 10 µl volume, using 400 units per reaction of T4 DNA Ligase enzyme from New England Biolabs (catalog #202L), with the accompanying reaction buffer. Ligation reactions were done at room temperature for 1 hour, with 1 µl of each of the Qiagen-purified fragments (approximately 20-50 ng of each DNA, depending on recovery from the gel isolation). Ligation products were then transformed into *E. coli* strain JM 109 and plated onto an appropriate growth and selection medium, such as LB with 100µg/ml of ampicillin to select for transformants.

For each ligation reaction, six transformants were tested to determine if the desired construct was present. Plasmid DNA was purified and isolated using the QiaPrep Spin Mini Prep Kit, according to manufacturer's instructions. The constructs were verified by DNA sequencing and by restriction mapping.

Expression and purification of the chimeric enzymes was done as follows. Plasmids were transformed into *E. coli* strain JM109 (Promega). Log phase cultures (200 ml) of JM109 were induced with 0.5 mM IPTG (Promega) and grown for an additional 16 hours prior to harvest. Crude extracts containing soluble proteins were prepared by lysis of pelleted cells in 5 ml of 10 mM Tris-HCl, pH 8.3, 1mM EDTA, 0.5mg/ml lysozyme during incubation at room temperature for 15 minutes. The lysate was mixed with 5 ml of 10 mM Tris-HCl pH 7.8, 50 mM KCl, 1 mM EDTA, 0.5% Tween 20, 0.5% Nonidet P-40, heated at 72°C for 30 minutes, and cell debris was removed by centrifugation at 12,000x g for 5 minutes. Final purification of the protein was done by affinity chromatograpy using an Econo-Pac heparin cartridge (Bio-Rad) and Dionex DX 500 HPLC instrument. Briefly, the cartridge was equilibrated with 50mM Tris-HCl pH 8, 1 mM EDTA, and an enzyme extract dialyzed against the same buffer was loaded on the column and eluted with a linear gradient of NaCl (0-2 M) in the same buffer. The HPLC-purified protein was dialyzed and stored in 50% (vol/vol) glycerol, 20 mM Tris-HCl pH 8.0, 50 mM KCl, 0.5% Tween 20, 0.5% Nonidet P-40, and 100µg/m BSA. The enzymes were purified to homogeneity according to SDS-PAGE, and the enzyme concentrations were determined by measuring absorption at 279 nm.

1. Construction of TaqTth(N) and TthTaq(N)

The first exchange that was performed involved the polymerase domains of the two enzymes. Separation of the nuclease domain (the N-terminal end of the protein) from the polymerase domain (the C-terminal portion of the protein) was accomplished by cutting both genes with the restriction endonucleases EcoRI and NotI. The approximately 900 base pair fragment from the Tth DN RX HT gene was cloned into the homologous sites of the Taq DN RX HT gene, and the approximately 900 base pair fragment from the Taq DN RX HT gene was cloned into the homologous sites of the Tth DN RX HT gene, yielding two chimeras, TaqTth(N) (DNA sequence SEQ ID NO:74; amino acid sequence SEQ ID NO:75) which has the Taq DN RX HT 5' nuclease domain and the Tth DN RX HT polymerase domain, and TthTaq(N) (DNA sequence SEQ ID NO:76; amino acid sequence SEQ ID NO:77) which is made up of the Tth DN RX HT 5' nuclease domain and the Taq DN RX HT polymerase domain.

2. Construction of TaqTth(N-B)

The Taq DN RX HT construct was cut with the enzymes NdeI and BamHI and the larger, vector fragment was gel isolated as detailed above. The Tth DN RX HT construct was also cut with NdeI and BamHI and the smaller (approximately 795 base pairs) Tth fragment was gel isolated and purified. The Tth NdeI-BamHI insert was ligated into the Taq NdeI-BamHI vector as detailed above to generate the TaqTth(N-B) (DNA sequence SEQ ID NO:78; amino acid sequence SEQ ID NO:79).

3. Construction of TaqTth(B-S)

The Taq DN RX HT construct was cut with the enzymes BamHI and SalI and the larger vector fragment was gel isolated as detailed above. The Tth DN RX HT construct was also cut with BamHI and SalI and the smaller (approximately 741 base pairs) Tth fragment was gel isolated and purified. The Tth BamHI-SalI insert was ligated into the Taq BamHI-SalI vector as detailed above to generate the TaqTth(B-S) (DNA sequence SEQ ID NO:80; amino acid sequence SEQ ID NO:81).

4. Construction of TaqTth(N-D)

The Taq DN RX HT construct was cut with the enzymes NotI and NdeI and the larger vector fragment was isolated as detailed above. The Tth DN RX HT construct was also cut with NotI and NdeI and the smaller (approximately 345 base pairs) Tth fragment was gel isolated and purified. The Tth NotI-NdeI insert was ligated into the Taq NotI-NdeI vector as detailed above to generate the TaqTth(N-D) (DNA sequence SEQ ID NO:82; amino acid sequence SEQ ID NO:83).

5. Construction of TaqTth(D-B)

The Taq DN RX HT construct was cut with the enzymes NdeI and BamHI and the larger vector fragment was isolated as detailed above. The Tth DN RX HT construct was also cut with NdeI and BamHI and the smaller (approximately 450 base pairs) Tth fragment was gel isolated and purified. The Tth NdeI-BamHI insert was ligated into the Taq NdeI-BamHI vector as detailed above to generate the TaqTth(D-B) (DNA sequence SEQ ID NO:84; amino acid sequence SEQ ID NO:85).

6. Construction of TaqTth(Bs-B)

The Taq DN RX HT construct was cut with the enzymes BstBI and BamHI and the larger vector fragment was isolated as detailed above. The Tth DN RX HT construct was also cut with BstBI and BamHI and the smaller (approximately 633 base pairs) Tth fragment was gel isolated and purified. The Tth NdeI-BamHI insert was ligated into the Taq NdeI-BamHI vector as detailed above to generate TaqTth(Bs-B) (DNA sequence SEQ ID NO:86; amino acid sequence SEQ ID NO:87).

7. Construction of TaqTth(N-Bs)

The Taq DN RX HT construct was cut with the enzymes NotI and BstBI and the larger vector fragment was isolated as detailed above. The Tth DN RX HT construct was also cut with NotI and BstBI and the smaller (approximately 162 base pairs) Tth fragment was gel isolated and purified. The Tth NotI-BstBI insert was ligated into the Taq NotI-BstBI vector as detailed above to generate TaqTth(N-Bs) (DNA sequence SEQ ID NO:88; amino acid

sequence SEQ ID NO:89).

8. Construction of TthTaq(B-S)

The Tth DN RX HT construct was cut with the enzymes BamHI and SalI and the larger vector fragment was isolated as detailed above. The Taq DN RX HT construct was also cut with BamHI and SalI and the smaller (approximately 741 base pairs) Tth fragment was gel isolated and purified. The Taq BamHI-SalI insert was ligated into the Tth BamHI-SalI vector as detailed above to generate the TthTaq(B-S) (DNA sequence SEQ ID NO:90; amino acid sequence SEQ ID NO:91).

9. Construction of Tth Taq(N-B)

The Tth DN RX HT construct was cut with the enzymes NotI and BamHI and the larger vector fragment was isolated as detailed above. The Taq DN RX HT construct was also cut with NotI and BamHI and the smaller (approximately 795 base pairs) Tth fragment was gel isolated and purified. The Taq NotI-BamHI insert was ligated into the Tth NotI-BamHI vector as detailed above to generate the TthTaq(N-B) (DNA sequence SEQ ID NO:92; amino acid sequence SEQ ID NO:93).

The cleavage activities of these chimerical proteins were characterized as describe in Example 1, part A, and a comparison of the cleavage cycling rates on an RNA target is shown in Figure 12. As further discussed in the Description of the Invention, these data show that elements found in the central third of the TthPol protein are important in conferring the TthPol-like RNA-dependent cleavage activity on the chimerical proteins comprising portions of TaqPol.

EXAMPLE 4

Alterations influencing RNA-dependent 5' nuclease activity do not necessarily influence RNA-dependent DNA polymerase activity

TthPol is known to have a more active RNA template dependent DNA polymerase

than does the TaqPol (Myers and Gelfand, Biochemistry 30:7661 [1991]). To determine whether the RNA template dependent 5' nuclease activity of the Thermus DNA Pol I enzymes is related to their RNA-dependent polymerase activity, the D785N and D787N mutations used to create the polymerase-deficient versions of TaqPol and TthPol, respectively were reversed. Polymerase activity was similarly restored to the TaqTth (N) (DNA sequence SEQ ID NO:94; amino acid sequence SEQ ID NO:95), TaqTth(N-B) (DNA sequence SEQ ID NO:96; amino acid sequence SEQ ID NO:97), TaqTth(B-S) (DNA sequence SEQ ID NO:98; amino acid sequence SEQ ID NO:99) chimeras, and the TaqPol(W417L/G418K/ E507Q) (DNA sequence SEQ ID NO:100; amino acid sequence SEQ ID NO:101) mutant proteins.

Polymerase function was restored in all the above mentioned enzyme mutants by inserting the BamHI to Sall fragment of the native, non-DN sequence into the selected chimera or mutant enzyme. For example, the mutant construct TaqTth(N-B) was cut with the restriction enzyme BamHI (approximate amino acid position 593) and the restriction enzyme Sall (approximate amino acid position 840). The larger vector fragment was gel purified as described in Example 3D. The native TaqPol construct was also cut with the restriction endonucleases BamHI and Sall, and the smaller insert fragment containing the native amino acid sequence was also gel purified. The insert fragment was then ligated into the vector as detailed in Experimental Example 3D.

The polymerase activities of these proteins were evaluated by extension of the dT₂₅₋₃₅-oligonucleotide primer with fluorescein-labeled dUTP in the presence of either poly(dA) or poly(A) template. Primer extension reactions were carried out in 10 µl buffer containing 10 mM MOPS, pH7.5, 5 mM MgSO₄, 100 mM KCl. Forty ng of enzyme were used to extend 10 µM (dT)₂₅₋₃₀ primer in the presence of either 10 µM poly(A)₂₈₆ or 1 µM poly(dA)₂₇₃ template, 45 µM dTTP and 5 µM FI-dUTP at 60°C for 30 min. Reactions were stopped with 10 µl of stop solution (95% formamide, 10 mM EDTA, 0.02% methyl violet dye). Samples (3 µl) were fractionated on a 15% denaturing acrylamide gel and the fraction of incorporated FI-dUTP was quantitated using an FMBIO-100 fluorescent gel scanner (Hitachi) equipped with a 505 nm filter as described above.

As shown in Figure 16, the DNA-dependent polymerase activities are very similar for all constructs used in this experiment, whereas the RNA-dependent polymerase activities of

TthPol, TaqTth(N) and TaqTth(B-S) are at least 6-fold higher than the activities of TaqPol, TaqTth(N-B) and the TaqPol W417L/G418K/E507Q mutant. From the analysis of these results, it can be concluded that the high RNA-dependent DNA polymerase activity of TthPol is determined by the C-terminal half of the polymerase domain (roughly, amino acids 593-830) and that the RNA-dependent 5' nuclease and polymerase activities are not related to each other, and are controlled by different regions.

EXAMPLE 5

Specific point mutants in Taq DN RX HT developed from information from the chimeric studies

The chimeric studies (Example 3, above) suggest that the part of the TthPol sequence determining its high RNA-dependent 5' nuclease activity comprises the BstBI-BamHI region located approximately between amino acid 382 and 593. Comparison of the amino acid sequences between the BstBI and BamHI regions of Tth DN RX HT and Taq DN RX HT (SEQ ID NOS:20 and 19, respectively) revealed only 25 differences (Figure 13). Among these, 12 amino acid changes were conservative while 13 of the differences resulted in a changes in charge. Since the analysis of the chimeric enzymes suggested that the critical mutations are located in both the BstBI-NdeI and the NdeI-BamHI regions of Tth DN RX HT, site specific mutagenesis was used to introduce the Tth DN RX HT specific amino acids into the BstBI-NdeI and NdeI-BamHI regions of the TaqTth(D-B) and the TaqTth(N-D) respectively.

Six Tth DN RX HT specific substitutions were generated in the BstBI-NdeI region of the TaqTth(D-B) by single or double amino acid mutagenesis. Similarly, 12 Tth DN RX HT specific amino acid changes were introduced at the homologous position of the NdeI-BamHI region of the TaqTth(N-D).

Plasmid DNA was purified from 200 ml of JM109 overnight culture using QIAGEN Plasmid Maxi Kit (QIAGEN, Chatsworth, CA) according to the manufacturer's protocol to obtain enough starting material for all mutagenesis reactions. All site specific mutations were introduced using the Transformer Site Directed mutagenesis Kit (Clontech) according to the

manufacturer's protocol; specific sequence information for the mutagenic primers used for each site is provided below. One of two different selection primers, Trans Oligo AlwNI/SpeI or Switch Oligo SpeI/AlwNI (Clontech, Palo Alto, CA catalog #6488-1 or catalog #6373-1) was used for all mutagenesis reactions described. The selection oligo used in a given reaction is dependent on the restriction site present in the vector. All mutagenic primers were synthesized by standard synthetic chemistry. Resultant colonies were *E.coli* strain JM109.

1. Construction of TaqTth(D-B) E404H (DNA sequence SEQ ID NO:102; amino acid sequence SEQ ID NO:103)

Site specific mutagenesis was performed on pTrc99A TaqTth(D-B) DNA using the mutagenic primer 240-60-01 5'-gag gag gcg ggg cac cgg gcc gcc ctt-3' (SEQ ID NO:104) to introduce the E404H mutation.

2. Construction of TaqTth(D-B) F413H/A414R (DNA sequence SEQ ID NO:105; amino acid sequence SEQ ID NO:106)

Site specific mutagenesis was performed on pTrc99A TaqTth(D-B) DNA using the mutagenic primer 240-60-02 5'-ctt tcc gag agg ctc cat cgg aac ctg tgg ggg agg-3' (SEQ ID NO:107) to introduce the F413H and the A414R mutations.

3. Construction of TaqTth(D-B) W417L/G418K (DNA sequence SEQ ID NO:108; amino acid sequence SEQ ID NO:109)

Site specific mutagenesis was performed on pTrc99A TaqTth(D-B) DNA using the mutagenic primer 240-60-03 5'-ctc ttc gcc aac ctg ctt aag agg ctt gag ggg gag-3' (SEQ ID NO:110) to introduce the W417L and the G418K mutations.

4. Construction of TaqTth(D-B) A439R (DNA sequence SEQ ID NO:111; amino acid sequence SEQ ID NO:112)

Site specific mutagenesis was performed on pTrc99A TaqTth(ND-B) DNA using the mutagenic primer 240-60-04 5'-agg ccc ctt tcc cgg gtc ctg gcc cat-3' (SEQ ID NO:113) to introduce the A439R mutation.

**5. Construction of TaqTth(N-D) L451R (DNA sequence SEQ ID NO:114;
amino acid sequence SEQ ID NO:115)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-05 5'-acg ggg gtg cgc cgg gac gtg gcc tat-3' (SEQ ID NO:116) to
introduce the L415 mutation.

**6. Construction of TaqTth(N-D) R457Q (DNA sequence SEQ ID NO:117;
amino acid sequence SEQ ID NO:118)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-06 5'-gtg gcc tat ctc cag gcc ttg tcc ctg-3' (SEQ ID NO:119) to
introduce the L415Q mutation.

**7. Construction of TaqTth(N-D) V463L (DNA sequence SEQ ID NO:120;
amino acid sequence SEQ ID NO:121)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-07 5'-ttg tcc ctg gag ctt gcc gag gag atc-3' (SEQ ID NO:122) to
introduce the V463L mutation.

**8. Construction of TaqTth(N-D) A468R (DNA sequence SEQ ID NO:123;
amino acid sequence SEQ ID NO:124)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-08 5'-gcc gag gag atc cgc cgc ctc gag gcc-3' (SEQ ID NO:125) to
introduce the A468R mutation.

**9. Construction of TaqTth(N-D) A472E (DNA sequence SEQ ID NO:126;
amino acid sequence SEQ ID NO:127)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-09 5'-gcc cgc ctc gag gag gag gtc ttc cgc-3' (SEQ ID NO:128) to
introduce the A472E mutation.

**10. Construction of TaqTth(N-D) G499R (DNA sequence SEQ ID NO:129;
amino acid sequence SEQ ID NO:130)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-10 5'-ttt gac gag cta agg ctt ccc gcc atc-3' (SEQ ID NO:131) to
introduce the G499R mutation.

**11. Construction of TaqTth(N-D) E507Q (DNA sequence SEQ ID NO:132;
amino acid sequence SEQ ID NO:133)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 276-046-04 5'-atc gcc aag acg caa aag acc ggc aag-3' (SEQ ID NO:134) to
introduce the E507Q mutation.

**12. Construction of TaqTth(N-D) Y535H (DNA sequence SEQ ID NO:135;
amino acid sequence SEQ ID NO:136)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-11 5'-aag atc ctg cag cac cgg gag etc acc-3' (SEQ ID NO:137) to
introduce the Y535H mutation.

**13. Construction of TaqTth(N-D) S543N (DNA sequence SEQ ID NO:138;
amino acid sequence SEQ ID NO:139)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-12 5'-acc aag ctg aag aac acc tac att gac-3' (SEQ ID NO:140) to
introduce the S543N mutation.

**14. Construction of TaqTth(N-D) I546V (DNA sequence SEQ ID NO:141;
amino acid sequence SEQ ID NO:142)**

Site specific mutagenesis was performed on pTrc99AtaqTth(N-D) DNA using the
mutagenic primer 240-60-13 5'-aag agc acc tac gtg gac ccc ttg ccg-3' (SEQ ID NO:143) to
introduce the I546V mutation.

15. Construction of TaqTth(N-D) D551S/I553V (DNA sequence SEQ ID NO:144; amino acid sequence SEQ ID NO:145)

Site specific mutagenesis was performed on pTrc99A_{TaqTth}(N-D) DNA using the mutagenic primer 240-60-14 5'-att gac ccc ttg ccg agc ctc gtc cac ccc agg acg ggc-3' (SEQ ID NO:146) to introduce the D551S and the I553V mutations.

16. Construction of TaqDN RX HT W417L/G418K/E507Q (DNA sequence SEQ ID NO:147; amino acid sequence SEQ ID NO:148)

The TaqDN RX HT W417L/G418K/E507Q triple mutant was made by combining the TaqTth(D-B)W417L/G418K with the TaqTth(N-D) E507Q. TaqTth(D-B)W417L/G418K was cut with the restriction enzymes NdeI and BamHI, and the larger, vector fragment was isolated as detailed in Example 3. The TaqTth(N-D) E507Q construct was also cut with NdeI and BamHI and the smaller (approximately 795 base pairs) fragment was gel isolated and purified as detailed in Example 3. The NdeI-BamHI insert was ligated into the gel purified vector, as detailed in Example 3.

17. Construction of TaqDN RX HT W417L/E507Q (DNA sequence SEQ ID NO:149; amino acid sequence SEQ ID NO:150)

Starting with TaqDN RX HT W417L/G418K/E507Q described above, mutagenic primer 337-01-02: 5'-TTC GCC AAC CTG CTT GGG AGG CTT GAG GGG GAG -3' (SEQ ID NO:151) was used in a site specific mutagenesis reaction to change the K at amino acid position 418 back to the wild-type amino acid, G. Site specific mutagenesis was done using the Transformer Site Directed Mutagenesis Kit (Clonetech) according to the manufacturer's instructions, and as described in Experimental Example 4.

18. Construction of TaqDN RX HT G418K/E507Q (DNA sequence SEQ ID NO:152; amino acid sequence SEQ ID NO:153)

Starting with TaqDN RX HT W417L/G418K/E507Q described above, mutagenic primer 337-01-01: 5'-CTC TTC GCC AAC CTG TGG AAG AGG CTT GAG GGG -3' was

used in a site specific mutagenesis reaction to change the L at amino acid position 417 back to the wild-type amino acid, W. Site specific mutagenesis was done using the Transformer Site Directed Mutagenesis Kit (Clontech) according to the manufacturer's instructions, and as described in Experimental Example 4.

Expression and purification of mutant proteins was done as detailed in Example 3, and the cleavage activities of these proteins were characterized as describe in Example 1, part A. A comparison of the cleavage cycling rates of a selection of these mutant proteins on an RNA target is shown in Figure 14. As further discussed in the Description of the Invention, these data show that amino acids in the regions 417/418 and amino acid 507 are important in the conferring the TthPol-like RNA-dependent cleavage activity on the chimerical proteins comprising portions of TaqPol in combination with portions of TthPol that are not independently capable of providing enhanced RNA dependent activity (*i.e.*, the D-B and N-D portions of Tth). As described in the Description of the Invention, Taq DN RX HT variant carrying only the W417L, G418K and E507Q substitutions were created. By comparing their cleavage rates to that of Tth DN RX HT on the IL-6 RNA substrate as described in Example 1, these mutations were determined to be sufficient to increase the Taq DN RX HT activity to the Tth DN RX HT level. Figure 15 shows that the Taq DN RX HT W417L/G418K/E507Q and Taq DN RX HT G418K/E507Q mutants have 1.4 times higher activity than Tth DN RX HT and more than 4 fold higher activity than Taq DN RX HT, whereas the Taq DN RX HT W417L/E507Q mutant has the same activity as the enzyme, which is about 3 fold higher than Taq DN RX HT. These results demonstrate that K418 and Q507 of TthPol are particularly important amino acids in providing RNA dependent 5' nuclease activity that is enhanced compared to TaqPol.

EXAMPLE 6

RNA-dependent 5' nuclease properties of the Taq DN RX HT G418K/E507Q 5' nuclease are similar to Tth DN RX HT with respect to salt and temperature optima

To determine if the G418K/E507Q mutations caused any significant changes in the properties of the Taq DN RX HT mutant in addition to the increased cleavage rate with the

RNA target, the Taq DN RX HT G418K/E507Q (SEQ ID NO:153), Taq DN RX HT (SEQ ID NO:73), and Tth DN RX HT (SEQ ID NO:72) enzymes were compared in the RNA template dependent 5' nuclease assay under conditions where temperature and the concentrations of salt and divalent ions were varied. The upstream DNA and the template RNA strands of the substrate used in this study were linked into a single IrT molecule (SEQ ID NO:27) as shown in Figure 20A, and the labeled downstream probe (SEQ ID NO:26) was present in large excess. The 5' end of the target RNA strand was blocked with a biotin-streptavidin complex to prevent any non-specific degradation by the enzyme during the reaction (Lyamichev *et al.*, Science 260:778 [1993], Johnson *et al.*, Science 269:238 [1995]).

The cleavage rates for Taq DN RX HT G418K/E507Q, Taq DN RX HT, and Tth DN RX HT are plotted as functions of temperature in Figure 20B. The closed circles represent enzyme Taq DN RX HT, the open circles represent enzyme Tth DN RX HT, and the Xs represent enzyme Taq DN RX HT G418K/E507Q. The difference in the activities of Tth and Taq DN RX HT enzymes with the IrT substrate is even greater than the difference found with the IL-6 RNA substrate when tested in a cleavage assay as described in Example 1. The G418K/E507Q mutations increase the activity of the Taq enzyme more than tenfold and by 25% compared with the Tth enzyme. All three enzymes show a typical temperature profile of the invasive signal amplification reaction and have the same optimal temperature. No significant effect of G418K/E507Q mutations on DNA dependent 5' nuclease activity of Taq DN RX HT with the all-DNA substrate analogous to IrT (SEQ ID NO:28) under the same conditions was found.

The effects of KCl and MgSO₄ concentrations on the 5' nuclease activity of Taq DN RX HT G418K/E507Q, Taq DN RX HT, and Tth DN RX HT with the IrT substrate are shown in Figure 20C and D. The activities of all enzymes have similar salt dependencies with an optimal KCl concentration of 100 mM for Taq DN RX HT G418K/E507Q and Tth DN RX HT and 50 mM for Taq DN RX HT. The optimal MgSO₄ concentration for all enzymes is approximately 8 mM. The analysis of the data presented in Figure 20 suggests that the properties of Taq DN RX HT G418K/E507Q are much closer to those of Tth DN RX HT rather than Taq DN RX HT confirming the key role of the G418K/E507Q mutations in the recognition of the substrate with an RNA target.

To understand the mechanism of the reduction of the 5' nuclease activity in the presence of an RNA versus a DNA target, the Michaelis constant (K_m) and the maximal catalytic rate (k_{cat}) of all three enzymes were determined, using an excess of the IrT substrate (SEQ ID NO:27) and the downstream probe (SEQ ID NO:26) and a limiting enzyme concentration. For these measurements, ten- μ l reactions were assembled containing 10 mM MOPS, pH 7.5, 0.05% Tween 20, 0.05% Nonidet P-40, 10 μ g/ml tRNA, 4 mM $MgCl_2$, 1 nM of enzyme (Taq DN RX HT, Tth DN RX HT, or Taq DN RX HT G418K/E507Q) and different concentrations (0.125, 0.25, 0.5 or 1 μ M) of an equimolar mixture of the IrT target and the downstream probe. The cleavage kinetics for each enzyme and each substrate concentration were measured at 46°C. Reactions were stopped by the addition of 10 μ l of 95% formamide containing 10 mM EDTA and 0.02% methyl violet (Sigma). One μ l of each stopped reaction digest was fractionated on a 20% denaturing acrylamide gel (19:1 cross-linked), with 7M urea, and in a buffer of 45 mM Tris-borate, pH 8.3, 1.4mM EDTA. Gels were scanned on an FMBIO-100 fluorescent gel scanner (Hitachi) using a 585 nm filter. The fraction of cleaved product (determined from intensities of bands corresponding to uncut and cut substrate with FMBIO Analysis software, version 6.0, Hitachi) was plotted as a function of reaction time. The initial cleavage rates were determined from the slopes of linear part of the cleavage kinetics and were defined as the concentration of cut product divided by the enzyme concentration and the time of the reaction (in minutes). The Michaelis constant K_m and the maximal catalytic rate k_{cat} of each enzyme with IrT substrate were determined from the plots of the initial cleavage rate as functions of the substrate concentration.

It was found that all three enzymes have similar K_m values (in the range of 200-300 nM) and k_{cat} values of approximately 4 min^{-1} for Taq DN RX HT and Tth DN RX HT and of 9 min^{-1} for Taq DN RX HT G418K/E507Q. That the G418K/E507Q mutations increase the k_{cat} of Taq DN RX HT more than two fold, but have little effect on K_m suggest that the mutations position the substrate in an orientation more appropriate for cleavage, rather than simply increase the binding constant.

EXAMPLE 7

Use of molecular modeling to further improve RNA-dependent 5' nuclease activity

A. Point mutants

To develop enzymes with altered function, sequence changes were introduced by site specific mutagenesis in predetermined locations or by random mutagenesis. Locations for site specific mutagenesis were chosen based on evidence from chimeric studies, relevant published literature, and molecular modeling. Seven additional mutant enzymes were developed from the Tth DN RX HT enzyme, and twenty additional mutant enzymes were developed from the Taq DN RX HT enzyme, both discussed previously. Some of the mutant enzymes are the result of multiple mutagenesis reactions, that is, more than one change has been introduced to obtain the final product. Mutation reactions were done using the Tth DN RX HT construct (SEQ ID NO:70) described in Example 2C2, or the Taq DN RX HT construct (SEQ ID NO:71), described in Example 2C1 unless otherwise stated. Plasmid DNA was purified from 200 ml of JM109 overnight culture using QIAGEN Plasmid Maxi Kit (QIAGEN, Chatsworth, CA) according to the manufacturer's protocol to obtain enough starting material for all mutagenesis reactions. All site specific mutations were introduced using the Transformer Site Directed mutagenesis Kit (Clontech) according to the manufacturer's protocol. One of two different selection primers, Trans Oligo AlwNI/SpeI or Switch Oligo SpeI/AlwNI (Clontech, Palo Alto CA catalog #6488-1 or catalog #6373-1) was used for all mutagenesis reactions described. The selection oligo used in a given reaction is dependent on the restriction site present in the vector. All mutagenic primers for both the site specific mutagenesis and the random mutagenesis were synthesized by standard synthetic chemistry. Resultant colonies for both types of reactions were *E.coli* strain JM109. Random mutagenesis methods are described below.

Mutants were tested via the rapid screening protocol detailed in Example 1. Then, if more detailed analysis was desired, or if a larger protein preparation was required, expression and purification of mutant proteins was done as detailed in Example 3.

1. Construction of Tth DN RX HT H641A, Tth DN RX HT H748A, Tth DN RX HT H786A

Site specific mutagenesis was performed on pTrc99A Tth DN RX HT DNA using the mutagenic primer 583-001-02: 5'-gct tgc ggt ctg ggt ggc gat gtc ctt ccc ctc-3' (SEQ ID NO:158) to introduce the H641A mutation (DNA sequence SEQ ID NO:156; amino acid sequence SEQ ID NO:157), or the mutagenic primer 583-001-03: 5' cat gtt gaa ggc cat ggc ctc cgc ggc ctc cct-3' (SEQ ID NO:161) to generate the H748A mutant (DNA sequence SEQ ID NO:159; amino acid sequence SEQ ID NO:160), or the mutagenic primer 583-001-04: 5'-cag gag gag ctc gtt ggc gac ctg gag gag-3' (SEQ ID NO:164) to generate the H786A mutant enzyme (DNA sequence SEQ ID NO:162; amino acid sequence SEQ ID NO:163).

2. Construction of Tth DN RX HT (H786A/G506K/Q509K)

Starting with the mutant Tth DN RX HT H786A, generated above, site specific mutagenesis was done using the mutagenic primer 604-022-02: 5'-gga ggc ctt gcc tgt ctt ctt cgt ctt ctt caa ggc ggg agg cct-3' (SEQ ID NO:167) to generate this variant termed "TthAKK", (DNA sequence SEQ ID NO:165; amino acid sequence SEQ ID NO:166).

3. Construction of Taq DN RX HT (W417L/G418K/E507Q/H784A)

Mutagenic oligonucleotide 158-029-02: 5'-gag gac cag ctc gtt ggc gac ctg aag gag cat-3' (SEQ ID NO:170) was used in a site specific mutagenesis reaction to introduce the H784A mutation and generate this construct termed "Taq4M" (DNA sequence SEQ ID NO:168; amino acid sequence SEQ ID NO:169).

4. Construction of Taq4M H639A, Taq4M R587A, Taq4M G504K and Taq4M G80E

Site specific mutagenesis was done on the Taq4M mutant, using primer 473-010-11: 5'-gaggggcgggacatcgccacggagaccgcccagc-3' (SEQ ID NO:173) to generate the Taq 4M H639A mutant (DNA sequence SEQ ID NO:171; amino acid sequence SEQ ID NO:172), primer 473-010-10: 5'-cag aac atc ccc gtc gcc acc ccg ctt ggg cag-3' (SEQ ID NO:176) to generate Taq 4M R587A (DNA sequence SEQ ID NO:174; amino acid sequence SEQ ID NO:175), primer 300-081-06: 5'-ggg ctt ccc gcc atc aag aag acg gag aag acc-3' (SEQ ID NO:179) to generate Taq 4M G504K (DNA sequence SEQ ID NO:177; amino acid sequence SEQ ID NO:178), and primer 330-088-04: 5'-cta ggg ctt ccc gcc atc aag aag acg caa aag acc ggc-3'

(SEQ ID NO:182) to generate the Taq 4M G80E mutant (DNA sequence SEQ ID NO:180; amino acid sequence SEQ ID NO:181).

5. Construction of Taq 4M P88E/P90E and Taq 4M L109F/A110T

Starting with Taq 4M described above, site specific mutagenesis was done using primer 473-087-03: 5'-ccg ggg aaa gtc ctc ctc cgt ctc ggc ccg gcc cgc ctt-3' (SEQ ID NO:185) to generate the P88E/P90E mutations (DNA sequence SEQ ID NO:183; amino acid sequence SEQ ID NO:184), or primer 473-087-05: 5'-cgg gac ctc gag gcg cgt gaa ccc cag gag gtc cac-3' (SEQ ID NO:188) to generate the L109F/A110T mutations (DNA sequence SEQ ID NO:186; amino acid sequence SEQ ID NO:187).

6. Construction of Taq DN RX HT (W417L/G418K/G499R/A502K/I503L/G504K/ E507K/H784A)

Two PCR reactions were performed, first using construct Taq4M (Taq W417L/G418K/G504K/E507Q/H784A) as a template. Using primers 158-84-01 5'-CTCCTCCACGAGTTCGGC-3' (SEQ ID NO:191) and 535-33-02 5'-ACC GGT CTT CTT CGT CTT CTT CAA CTT GGG AAG CCT GAG CTC GTC AAA-3' (SEQ ID NO:192) a 620 base pair PCR fragment was generated. Another 510 base pair PCR product was generated using primer 535-33-01 5'-AAG ACG AAG AAG ACC GGT AAG CGC TCC ACC AGC-3' (SEQ ID NO:193) and 330-06-03 5'-GTC GAC TCT AGA TCA GTG GTG GTG GTG GTG CTT GGC CGC CCG GCG CAT C-3' (SEQ ID NO:194). The two PCR products overlap such that a final recombinant PCR amplification was done using the outside primers 158-84-01 and 330-06-03 to yield the 1182 base pair product. The recombinant PCR product was digested with the restriction enzymes NotI and BamHI according to the manufacturer's instructions to yield a 793 base pair fragment. The parent plasmid Taq4M was also digested with the same enzymes and used as the vector for ligation. All DNA fragments were TAE agarose gel purified prior to ligation. The fragment was ligated into the vector, and transformed into JM109 cells, thus incorporating the mutations G499R, A502K, I503L, and E507K as well as the restriction endonuclease site, AgeI. This construct is termed "Taq 8M" (DNA sequence SEQ ID NO:189; amino acid sequence SEQ ID NO:190).

B. Random Mutagenesis

Numerous enzymes with altered function were generated via random mutagenesis. The regions of the protein targeted for random mutagenesis were chosen based on molecular modeling data and from information in the literature. Different mutagenic primers were used to introduce mutations into different regions of the protein. Random mutagenesis was performed on the Taq variant Taq 4M G504K (Taq DN RX HT W417L/G418K/G504K/E507Q/H784A/) (SEQ ID NO:177) described above and mutant PCR fragments generated in the mutagenesis reaction were exchanged for homologous regions in Taq8M (SEQ ID NO:189) unless otherwise stated.

Random mutagenesis was also performed on the Tth DN RX HT H786A (SEQ ID NO:162) described above. Mutant PCR fragments generated with the Tth DN RX HT H786A template were exchanged for homologous regions in the unaltered Tth DN RX HT H786A.

1. Random mutants in amino acid residues 500-507 or 513-520

The first mutagenic oligonucleotide, 535-054-01: 5'-gga gcg ctt acc ggt ctt (ttg cgt ctt ctt gat ctt ggg aag) cct tag ctc gtc aaa gag-3' (SEQ ID NO:195) was used in conjunction with 158-84-01: 5'-CTC CTC CAC GAG TTC GGC-3' (SEQ ID NO:196) to install random residues from amino acid position 500 to 507 of Taq polymerase variant Taq DN RX HT W417L/G418K/G504K/E507Q/H784A (SEQ ID NO:177). This was accomplished by synthesizing the primer 535-054-01 such that only 91% of the bases within the parenthesis are unaltered while the remaining 9% of the bases are an equal mixture of the other 3 nucleotides. The initial, unaltered sequence of this oligo includes the G499R, A502K and the Q507K changes.

To generate mutations in the region 500-507, primer 535-054-01 and primer 158-84-01 were used in a PCR reaction, using the Advantage cDNA PCR kit (Clontech) and Taq variant described above, as the target. This PCR fragments was then run on a 1% TAE agarose gel, excised and purified with QIAquick Gel Extraction Kit (Qiagen, Valencia CA, catalog # 28706). The purified fragment was cut with NotI and AgeI and ligated into pTaq8M that had been linearized with NotI and AgeI. JM109 *E.coli* cells (Promega) were transformed with the ligated products. Clones were tested as described below.

The second mutagenic oligonucleotide (used in a separate reaction) 535-054-02: 5'-caa aag acc ggt aag cgc (tcc acc agc gcc gcc gtc ctg gag) gcc ctc cgc gag gcc cac-3' (SEQ ID NO:197) was used in conjunction with 330-06-03: 5'-GTC GAC TCT AGA TCA GTG GTG GTG GTG GTG CTT GGC CGC CCG GCG CAT C-3' (SEQ ID NO:198) to install
5 random residues from amino acid 513-520. The bases within the parenthesis of primer 535-054-02 are also 91% wild-type and 3% each of the other 3 nucleotides.

To generate mutations in the region 513-520, primer 535-054-02 and primer 535-054-02 were used in a PCR reaction with Taq DN RX HT W417L/G418K/G504K/E507Q/H784A (SEQ ID NO:177) as template, as described above. The resulting PCR
10 fragment was purified as above and cut with the restriction enzymes AgeI and BamHI. The cut fragment was then ligated into the Taq8M construct, also linearized with AgeI and BamHI. JM109 *E.coli* cells were transformed with the ligated products. Clones were tested as described Example 1. Mutants developed from these include:

5 Taq DN RX HT W417L/G418K/G499R/A502K/K504N/E507K/H784A (M1-13) (DNA sequence SEQ ID NO:199; amino acid sequence SEQ ID NO:200).

Taq DN RX HT W417L/G418K/G499R/L500I/A502K/G504K/Q507H/H784A (M1-36) (DNA sequence SEQ ID NO:201; amino acid sequence SEQ ID NO:202).

20 Taq DN RX HT W417L/G418K/G499R/A502K/I503L/G504K/E507K/T514S/H784A (M2-24) (DNA sequence SEQ ID NO:203; amino acid sequence SEQ ID NO:204).

Taq DN RX HT W417L/G418K/G499R/A502K/I503L/G504K/E507K/ V518L/H784A (M2-06) (DNA sequence SEQ ID NO:205; amino acid sequence SEQ ID NO:206).

2. TthDN RX HT H786A random mutagenesis

To generate mutants in the helix-hairpin-helix region of the TthDN RX HT H786A
25 (SEQ ID NO:163) enzyme, two different PCR reactions were performed using the H786A (SEQ ID NO:162) mutant as a template. The two PCR products overlap such that a recombinant PCR reaction can be performed (Higuchi, in PCR Technology, H. A. Erlich, ed., Stockton Press, New York. pp61-70 [1989]). This final PCR product is then exchanged with the homologous region of the TthDN H786A mutant by using restriction enzyme sites located
30 on the ends of the fragment and within the TthDN H786A sequence.

Starting with TthDN H786A discussed above, and using primer 604-08-06: 5'-gtc gga ggg gtc ccc cac gag-3' (SEQ ID NO:207) and primer 390-76-08: 5'-tgt gga att gtg agc gg (SEQ ID NO:208), a 620 base pair PCR fragment was generated. PCR reactions were performed using the Advantage cDNA PCR kit (Clontech) according to manufacturer's instructions. This PCR product includes amino acids 1-194. No mutations were introduced via this reaction, however the restriction enzyme site EcoRI is present at the 5' end.

Starting with TthDN RX HT H786A discussed above, and using mutagenic primer 604-08-05: 5'-ctc gtg ggg gac ccc tcc gac aac ctc (ccc ggg gtc aag ggc atc ggg gag aag acc gcc) ctc aag ctt ctc aag-3' (SEQ ID NO:209) and primer 209-74-02: 5'-gtg gcc tcc ata tgg gcc agg ac-3' (SEQ ID NO:210) a 787 base pair PCR fragment was generated. PCR reactions were done as above. This fragment does contain random mutations, due to the presence of the mutagenic primer, 604-08-05. The bases within the parenthesis of this primer were synthesized such that 91% of the sequence is wild-type, while the additional 9% is evenly divided between the remaining 3 bases.

The two PCR fragments overlap, and were combined in a recombinant PCR reaction. Primers 390-76-08 and 209-74-02 were added, and the Advantage cDNA PCR kit (Clontech) was again used according to manufacturer's instructions. A 1380 base pair product was generated from this reaction.

The recombinant PCR product was cut with the restriction enzymes EcoRI and NotI according to the manufacturer's instructions to yield a 986 base pair fragment. TthDN RX HT H786A was prepared by cutting with the same enzymes. The fragment was then ligated into the vector, and transformed into JM109 cells. New mutants developed from this set of reactions include:

TthDN RX HT H786A/P197R/K200R (DNA sequence SEQ ID NO:211; amino acid sequence SEQ ID NO:212).

TthDN RX HT H786A/K205Y (DNA sequence SEQ ID NO:213; amino acid sequence SEQ ID NO:214).

TthDN RX HT H786A/G203R (DNA sequence SEQ ID NO:215; amino acid sequence SEQ ID NO:216).

3. Construction of Taq DN RX HT W417L/G418K/H784A

L109F/A110T/G499R/A502K/I503L/G504K/E507K/T514S (Taq SS)

Starting with Taq DN RX HT

W417L/G418K/G499R/A502K/I503L/G504K/E507K/T514S/H784A (SEQ ID NO:203) mutant described above, primer 473-087-05: 5'-cgg gac ctc gag gcg cgt gaa ccc cag gag gtc cac-3' (SEQ ID NO:219) was used in conjunction with the appropriate selection primer in a site specific mutagenesis reaction to incorporate the L109F and A110T mutations to generate this enzyme, termed "TaqSS" (DNA sequence SEQ ID NO:217; amino acid sequence SEQ ID NO:218).

4. Construction of Taq DN RX HT W417L/G418L/H784A

P88E/P90E/G499R/A502K/I503L/G504K/E507K/T514S

Starting with Taq DN RX HT

W417L/G418K/G499R/A502K/I503L/G504K/E507K/T514S/H784A (SEQ ID NO:203) mutant described above, primer 473-087-03: 5'-ccg ggg aaa gtc ctc ctc cgt ctc ggc ccg gcc cgc ctt-3' (SEQ ID NO:222) was used in conjunction with the appropriate selection primer in a site specific mutagenesis reaction to incorporate the P88E and P90E mutations to generate this enzyme (DNA sequence SEQ ID NO:220; amino acid sequence SEQ ID NO:221).

5. TaqSS random mutagenesis

Random mutagenesis was used to introduce additional changes in the helix-hairpin-helix domain of the TaqSS mutant (SEQ ID NO:217). The mutagenesis was done as described in example 9 above. In the first step, two different but overlapping PCR products were generated. One of the PCR products, generated with oligos 390-76-08 (SEQ ID NO:208), and 604-08-04: 5'-gtc gga ctc gtc acc ggt cag ggc-3' (SEQ ID NO:223) incorporates the EcoRI site into the fragment, but does not incorporate any mutations. The second PCR product utilizes mutagenic primer 604-08-03: 5'-ctg acc ggt gac gag tcc gac aac ctt (ccc ggg gtc aag ggc atc ggg gag aag acg gcg) agg aag ctt ctg gag-3' (SEQ ID NO:224) and primer 209-74-02 (SEQ ID NO:210). This fragment contains random point mutations, and when combined via recombinant PCR with the first fragment, can be cut with the restriction enzymes EcoRI and NotI, and ligated into the TaqSS construct, also cut with EcoRI and NotI. The ligated construct was then transformed into JM109. Colonies were

screened as described below. Enzymes developed from this mutagenesis include:

TaqSS K198N (DNA sequence SEQ ID NO:225; amino acid sequence SEQ ID NO:226).

TaqSS A205Q (DNA sequence SEQ ID NO:227; amino acid sequence SEQ ID NO:228).

5 TaqSS I200M/A205G (DNA sequence SEQ ID NO:229; amino acid sequence SEQ ID NO:230).

TaqSS K203N (DNA sequence SEQ ID NO:231; amino acid sequence SEQ ID NO:232).

TaqSS T204P (DNA sequence SEQ ID NO:233; amino acid sequence SEQ ID NO:234).

6. Construction of TaqSS R677A

10 To generate enzymes with sequence changes in both the arch region and in the polymerase region, additional specific point mutations were generated in TaqSS. Site specific mutagenesis was performed as described above using the oligo 473-060-10: 5'-tag ctc ctg gga gag ggc gtg ggc cga cat gcc-3' (SEQ ID NO:237) to generate the TaqSS R677A mutant (DNA sequence SEQ ID NO:235; amino acid sequence SEQ ID NO:236).

15 7. Construction of TaqTthAKK (DNA sequence SEQ ID NO:238; amino acid sequence SEQ ID NO:239) and TthTaq5M (DNA sequence SEQ ID NO:240; amino acid sequence SEQ ID NO:241)

20 Chimeric mutant TaqTthAKK and TthTaq5M were generated by cutting Tth DN RX HT (H786A/G506K/Q509K) (SEQ ID NO:165; here abbreviated TthAKK) or Taq 4M G504 (SEQ ID NO:177; here abbreviated Taq 5M) with the restriction endonucleases EcoRI and NotI. The smaller insert fragments as well as the larger vector fragments were gel purified as detailed in Example 3D, and the insert fragments were exchanged between the two mutants and ligated as described in Example 3D. Screening and verification of the construct sequence was also done as in Example 3D.

EXAMPLE 8

Improvement of RNA-dependent 5' nuclease activity in other polymerases

25 Information gained from the TaqPol/TthPol recombinations, mutagenesis and modeling, was used to make comparable mutations in additional DNA polymerases and

examined the effects on the cleavage activities of these enzymes. The DNA polymerases of *Thermus filiformus* (TfiPol) and *Thermus scotoductus* (TscPol) were cloned and purified as described in Example 2. The mutagenesis of these two proteins is described below.

A. Construction of TfiPolDN2M

Mutagenesis of pTrc99a-TfiPol (SEQ ID NO:48) was done using the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's protocol. The P420K mutation was made with the following two oligonucleotides; 5'-CTTCCAGAACCTCTTTAAACGGCTTCCGAGAAG (SEQ ID NO:244) and 5'-CTTCTCGGAAAGCCGTTTAAAGAGGTTCTGGAAG (SEQ ID NO:245). The E507Q mutation was made with the following two oligonucleotides; 5'-CCGGTGGGCCGGACGCAGAAGACGGGCAAGC (SEQ ID NO:246) and 5'-GCTTGCCCGTCTTCTGCGTCCGGCCCACCGG (SEQ ID NO:247). The D785N mutation was made with the following two oligonucleotides; 5'-CTCCTCCAAGTGCACAACGAGCTGGTCCTGG (SEQ ID NO:248) and 5'-CCAGGACCAGCTCGTTGTGCACTTGGAGGAG (SEQ ID NO:249). The plasmid containing all three mutations is called pTrc99a-TfiPolDN2M, (DNA sequence SEQ ID NO:242; amino acid sequence SEQ ID NO:243).

B. Construction of TscPolDN2M

Mutagenesis of pTrc99a-TscPol (SEQ ID NO:51) was done with the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's protocol. The E416K mutation was made with the following two oligonucleotides; 5'-GCCGCCCTCCTGAAGCGGCTTAAGGG (SEQ ID NO:252) and 5'-CCCTTAAGCCGCTTCAGGAGGGCGGC (SEQ ID NO:253). The E505Q mutation was made with the following two oligonucleotides; 5'-ATCGGCAAGACGCAGAAGACGGGCAAGC (SEQ ID NO:254) and 5'-GCTTGCCCGTCTTCTGCGTCTTGCCGAT (SEQ ID NO:255). The D783N mutation was made with the following two oligonucleotides; 5'-TTGCAGGTGCACAACGAAGTGGTCCTC (SEQ ID NO:256) and 5'-GAGGACCAGTTCGTTGTGCACCTGCAA (SEQ ID NO:257). The plasmid containing all three mutations is called pTrc99a-TscPolDN2M, (DNA sequence SEQ ID NO:250; amino

acid sequence SEQ ID NO:251).

C. Chimerics of Tsc, Tfi, Tth and Taq mutants

1. **Construction of TfiTth AKK (DNA sequence SEQ ID NO:258; amino acid sequence SEQ ID NO:259), TscTthAKK (DNA sequence SEQ ID NO:260; amino acid sequence SEQ ID NO:261), TfiTaq5M (DNA sequence SEQ ID NO:262; amino acid sequence SEQ ID NO:263) and TscTaq5M (DNA sequence SEQ ID NO:264; amino acid sequence SEQ ID NO:265)**

To generate chimeric enzymes between Tth DN RX HT (H86A/G506K/Q509K) (here abbreviated TthAKK, SEQ ID NO:165) or Taq 4M G504 (here abbreviated Taq 5M, SEQ ID NO:177), and Tfi DN 2M (SEQ ID NO:242), or Tsc DN 2M (SEQ ID NO:250), additional restriction endonuclease sites were introduced by site specific mutagenesis into the named Tfi and Tsc mutants. Mutagenic primers 700-011-01 5'-cag acc atg aat tcc acc cca ctt ttt gac ctg gag-3' (SEQ ID NO:275) and 700-011-02 5'-gtg gac gcg gcc gcc cga ggc cgc cgc cag ggc cag-3' (SEQ ID NO:276) were used to introduce an EcoRI site at amino acid position 1 and a NotI site at amino acid position 331 in Tfi DN 2M. Mutagenic primers 700-011-03 5'-cag acc atg aat tcc ctg ccc ctc ttt gag ccc aag-3' (SEQ ID NO:277) and 700-011-04 5'-gta aac cgc gcc gcc cca ggc ggc ggc caa ggc gtt-3' (SEQ ID NO:278) were used to introduce an EcoRI site at amino acid position 1 and a NotI site at amino acid position 327 in Tsc DN 2M. PCR reactions were done using the Advance cDNA PCR kit (Clontech) according to manufacturer's instructions and either Tfi DN 2M or Tsc DN 2M as target, with their corresponding primers. The 1017 base pair PCR products were cut with both EcoRI and NotI to yield 993 base pair insert fragments that were gel purified as described in Example 3D. The mutants Taq4M G504K (SEQ ID NO:177) and Tth DN RX HT (H786A/G506K/Q509K) (SEQ ID NO:165) were also cut with EcoRI and NotI, and the larger, vector fragment was gel isolated as above. Ligations were performed as detailed in Example 3D, as was the screening and verification of the new constructs.

TABLE 1

<u>Klenow</u>	<u>Kcat (s⁻¹)</u>	<u>Km(dNTP)</u> <u>(μM)</u>	<u>Kd (nM)</u>	<u>Relative DNA</u> <u>affinity</u>	<u>Reference</u>	<u>Taq Pol</u>
Wild-Type	2.4	2.8	8	1	2	Wild-Type
S610A	n. d.	-	n. d.	-	5	S515
R668A	0.006	6.5	140, 150	0.06, 0.05	1,2	R573*
N678A	n. d.	-	n. d.	-	5	N583
E710A	0.1	15	250	0.03	2	E615*
E710D	1.7	7.7	110	0.07	2	E615*
K758A	0.131	15.6	-	0.63	4	K663
K758R	2.0	2.1	-	1.125	4	K663
Y766S	0.8	6.4	13	0.4, 0.6	1,2	Y671
R841A	0.3	9.8	40, 53	0.2	1,2	R746*
N845A	1.0	23	8, 5	1.0, 1.7	1,2	N750*
N845Q	0.03	1.7	80, 55	0.1, 0.2	1,2	N750*
Q849A	0.02	3.8	100, 160	0.08, 0.05	1,2	Q754
Q849E	0.001	n. d.	90, 91	0.09	1,2	Q754
H881A	0.3	3.3	20, 28	0.4, 0.3	1,2	H784*
D882N	<0.0001	n. d.	30	0.6	2	D785
D882S	0.001	7.5	0.9	9	2	D785

References:

1. JBC (1990) 265:14579-14591
2. JBC (1992) 267:8417-8428
3. Eur. J. Biochem (1993) 214:59-65
4. JBC (1994) 269:13259-13265
5. Nature (1996) 382:278-281

TABLE 2: Rational mutations in the polymerase region**A. DNA activity table**

	IdT	%Tth	%Taq4M	HP	X
Tth DN RX HT	31.91	100%	83%	3.81	101.9
Tth DN RX HT H641A	23.61	74%	62%	5.32	221.24
Tth DN RX HT R748A	22.1	69%	58%	4.39	88.17
Tth DN RX HT H786A	34.31	108%	90%	7.75	185.35
Tth DN RX HT H786A/G506K/ Q509K (AKK)	32.1	101%	84%	5.7	332.8
Taq DN RX HT W417L/G418K/ E507Q/H784A (Taq 4M)	38.23	120%	100%	68.21	1100.18
Taq 4M G504K	36.04	113%	94%	31.76	417.40
Taq 4M H639A	42.95	135%	112%	91.46	2249.67
Taq 4M R587A	44.78	140%	117%	143.0	252.69
Taq DN RX HT W417L/G418K/ G499R/A502K/ I503L/K504N/ E07K/H784A (Taq8M)	43.95	138%	115%	122.53	346.56
TaqSS R677A	32.3	101%	84%	206.9	2450.0

B. RNA activity table

	IrT1	%Tth	%Taq4M
Tth DN RX HT	0.89	100%	34%
Tth DN RX HT H641A	1.18	133%	45%
Tth DN RX HT R748A	1.34	151%	51%
Tth DN RX HT H786A	1.31	147%	49%
Tth DN RX HT H786A/G506K/ Q509K (AKK)	1.59	179%	60%
Taq DN RX HT W417L/G418K/ E507Q/H784A (Taq 4M)	2.65	298%	100%
Taq 4M G504K	2.76	310%	114%
Taq 4M H639A	3.89	437%	147%
Taq 4M R587A	3.13	352%	118%
Taq DN RX HT W417L/G418K/ G499R/A502K/ I503L/K504N/ E07K/H784A (Taq8M)	4.00	450%	151%
TaqSS R677A	2.22	249%	84%

TABLE 3: Rational arch mutations**DNA activity table**

	IdT	%Tth	%Taq4M	HP	X
Taq 4M P88E/P90E	10.20	32%	27%	2.00	97.00
Taq 4M G80E	26.30	82%	69%	103.6	2900
Taq 4M L109F/A110T	36.45	114%	95%	19.71	749.69

RNA activity table

	IrT1	%Tth	%Taq4M		
Taq 4M P88E/P90E	Taq 4M P88E/P90E	0.10	11%	4%	
Taq 4M G80E	Taq 4M G80E	3.11	349%	117%	
Taq 4M L109F/A110T	Taq 4M L109F/A110T	2.45	275%	92%	

TABLE 4: Arch/thumb combinations**DNA activity table**

	IdT	%Tth	%Taq4M	HP	X
Taq W417L/ G418K/E507K/ H784A/L109F/ A110T/G499R/ A502K/I503L/ G504K/E507K/ T514S (Taq SS)	63.33	198%	166%	177.05	202.32
Taq P88E/P90E/ W417L/G418K/ G499R/A502K/ I503L/G504K/ E507K/T514S/ H784A	36.48	114%	95%	9.44	70.35

RNA activity table

	IrT1	%Tth	%Taq4M
Taq W417L/ G418K/E507K/ H784A/L109F/ A110T/G499R/ A502K/I503L/ G504K/E507K/ T514S (Taq SS)	3.16	355%	119%
Taq P88E/P90E/ W417L/G418K/ G499R/A502K/ I503L/G504K/ E507K/T514S/ H784A	0.22	25%	8%

TABLE 5: Helix-hairpin-helix random mutagenesis**DNA activity table**

	IdT	%Tth	%Taq4M	HP	X
TaqSS K198N	23.4	73%	61%	25.7	1233.1
TaqSS A205Q	25.6	80%	67%	13.4	699.1
TaqSS T204P	11.2	35%	29%	1.9	209.4
TaqSS I200M/A205G	16.8	53%	44%	7.8	597.2
TaqSS K203N	25.9	81%	68%	36.6	1429.8
Tth DN RX HT H786A/P197R/K200R	10.7	33%	28%	3.2	66.3
Tth DN RX HT H786A/K205Y	11.5	36%	30%	6.1	327.5
Tth DN RX HT H786A/G203R	18.3	57%	48%	2.1	98.8

RNA activity table

	IrT1	%Tth	%Taq4M
TaqSS K198N	1.22	137%	46%
TaqSS A205Q	0.62	70%	23%
TaqSS T204P	0.36	40%	14%
TaqSS I200M/A205G	0.77	87%	29%
TaqSS K203N	2.09	235%	79%
Tth DN RX HT H786A/P197R/K200R	0.47	52%	18%
Tth DN RX HT H786A/K205Y	0.68	77%	26%
Tth DN RX HT H786A/G203R	1.61	180%	61%

TABLE 6: Random thumb mutations
DNA activity table

	IdT	%Tth	%Taq4M	HP	X
Taq DN RX HT W417L/G418K/ E507K/H784A/G499R/ A502K/K504N/ (M1-13)	59.96	188%	157%	133.65	907.41
Taq DN RX HT W417L/G418K/ /H784A/L500I/Q507H A502K/G504K (M1-36)	46.74	146%	122%	123.11	822.61
Taq DN RX HT W417L/G418K/G499R/ A502K/G504K/E507K/ H784A/T514S (M2-24)	85.7	269%	224%	369.96	3752.12
Taq DN RX HT W417L/G418K/G499R/ A502K/G504K/E507K/ H784A/V518L (M2-06)	76.7	240%	201%	355.87	2038.19

RNA activity table

	IrT1	%Tth	%Taq4M
Taq DN RX HT W417L/G418K/ E507K/H784A/G499R/ A502K/K504N/ (M1-13)	2.55	287%	96%
Taq DN RX HT W417L/G418K/ /H784A/L500I/Q507H A502K/G504K (M1-36)	2.71	304%	102%
Taq DN RX HT W417L/G418K/G499R/ A502K/G504K/E507K/ H784A/T514S (M2-24)	4.43	498%	167%
Taq DN RX HT W417L/G418K/G499R/ A502K/G504K/E507K/ H784A/V518L (M2-06)	3.56	400%	134%

TABLE 7: Chimeric mutants**A. DNA activity table**

	IdT2	%TthAKK	HP	X
TthAKK	34.18	100%	5	393
Taq 4M G504K	40.19	105%	28	1991
Tfi DN 2M	36.60	106%	289	1326
Tsc DN 2M	25.49	75%	283	2573
TaqTthAKK	63.89	187%	32	1658
TthTaq 4M G504K	25.03	73%	8	627
TfiTthAKK	34.13	100%	15	459
TscTthAKK	35.23	103%	29	2703
TfiTaq 4M G504K	35.69	104%	37	872
TscTaq 4M G504K	30.04	88%	25	2008

B. RNA activity table

	IrT3	%TthAKK
TthAKK	2.27	100%
Taq 4M G504K	2.31	102%
Tfi DN 2M	0.20	9%
Tsc DN 2M	0.29	13%
TaqTthAKK	6.81	300%
TthTaq 4M G504K	1.09	48%
TfiTthAKK	1.24	55%
TscTthAKK	9.65	4.25%
TfiTaq 4M G504K	1.05	46%
TscTaq 4M G504K	2.95	130%

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in the relevant fields are intended to be within the scope of the following claims.

5

SEQUENCE LISTING

<110> Ma, Wu-Po
Lyamichev, Victor I.
Kaiser, Michael W.
Lyamicheva, Natalie E.
Allawi, Hatim T.
Schaefer, James J.
Neri, Bruce P.

<120> Improved Enzymes for the Detection of Specific Nucleic Acid Sequences

<130> FORS-04323

<140> not yet assigned

<141> 2000-05-24

<160> 278

<170> PatentIn Ver. 2.0

<210> 1

<211> 2506

<212> DNA

<213> Thermus aquaticus

<400> 1

```
atgaggggga tgctgcccct ctttgagccc aagggccggg tcctcctggt ggacggccac 60
cacctggcct accgcacctt ccacgccctg aagggcctca ccaccagccg gggggagccg 120
gtgcaggcgg tctacggctt cgccaagagc ctctcaagg ccctcaagga ggacggggac 180
gcggtgatcg tgggtcttga cgccaaggcc ccctccttcc gccacgaggc ctacgggggg 240
tacaaggcgg gccgggcccc cacgccggag gactttcccc ggcaactcgc cctcatcaag 300
gagctggttg acctcctggg gctggcgcg ctcgaggtcc cgggctacga ggcgagac 360
gtcctggcca gcctggccaa gaaggcggaa aaggagggt acgaggtccg catcctcacc 420
gccgacaaag acctttacca gctcctttcc gaccgcatcc acgtcctcca ccccgagggg 480
tacctcatca ccccggcctg gctttgggaa aagtagggcc tgaggcccga ccagtgggcc 540
gactaccggg ccctgaccgg ggacgagtc gacaaccttc ccgggggtcaa gggcatcggg 600
gagaagacgg cgaggaagct tctggaggag tgggggagcc tggaagccct cctcaagaac 660
ctggaccggc tgaagcccg catccgggag aagatcctgg ccacatgga cgatctgaag 720
ctctcctggg acctggccaa ggtgcgcacc gacctgcccc tggaggtgga ctctgccaaa 780
aggcgggagc ccgaccggga gaggcttagg gcctttctgg agaggcttga gtttggcagc 840
ctcctccacg agttcggcct tctggaaaag cccaaggccc tggaggaggc cccctggccc 900
ccgccggaag gggccttcgt gggctttgtg ttttcccgca aggagcccat gtgggccgat 960
cttctggccc tggccgccc cagggggggc cgggtccacc gggccccga gccttataaa 1020
gccctcaggg acctgaagga ggcgcggggg cttctcgcca aagacctgag cgttctggcc 1080
```


aaggcggggc gggcccccac cccggaggac tttccccggc agctggccct catcaaggag 300
ttggtggacc tcctaggcct tgtgcggctg gaggttcccg gctttgaggc ggacgacgtg 360
ctggccaccc tggccaagcg ggcggaaaag gaggggtacg aggtgcgcat cctcactgcc 420
gaccgcgacc tctaccagct cctttcggag cgcacgcga tcctccaccc tgaggggtac 480
ctgatcacc cggcgtggct ttacgagaag tacggcctgc gcccgagca gtgggtggac 540
taccggggcc tggcggggga cccctcggat aacatccccg ggggaaggc catcggggag 600
aagaccgccc agaggctcat ccgcgagtg gggagcctgg aaaacctctt ccagcacctg 660
gaccaggtga agcctcctt gcgggagaag ctccaggcgg gcatggaggc cctggccctt 720
tcccgaagc tttcccaggt gcacactgac ctgcccctgg aggtggactt cgggaggcgc 780
cgcacacca acctggaggg tctgcgggct tttttggagc ggttggagtt tggaagcctc 840
ctccacgagt tcggcctcct ggagggggcg aaggcggcag aggaggcccc ctggccccct 900
ccggaagggg cttttttggg cttttccttt tcccgtcccg agcccatgtg ggccgagctt 960
ctggccctgg ctggggcgct ggagggggcg ctccatcggg cacaagacc ccttaggggc 1020
ctgagggacc ttaaggggt gcggggaatc ctggccaagg acctggcggt tttggccctg 1080
cgggagggcc tggacctctt cccagaggac gaccccatgc tcctggccta ccttctggac 1140
ccctccaaca ccacctga ggggggtggc cggcgttacg ggggggagtg gacggaggat 1200
gcgggggaga gggccctcct ggccgagcgc ctcttcaga ccctaaagga gcgccttaag 1260
ggagaagaac gcctgctttg gctttacgag gaggtggaga agccgctttc ccgggtgttg 1320
gcccgatgg aggccacggg ggtccggctg gacgtggcct acctccaggc cctctccctg 1380
gaggtggagg cggaggtgcg ccagctggag gaggaggtct tccgcctggc cggccacccc 1440
ttcaacctca actcccgga ccagctggag cgggtgctct ttgacgagct gggcctgcct 1500
gccatcggca agacggagaa gacggggaaa cgctccacca gcgctgccgt gctggaggcc 1560
ctgcgagagg cccaccccat cgtggaccgc atcctgcagt accgggagct caccaagctc 1620
aagaacacct acatagacc cctgcccgc ctggtccacc ccaagaccg ccggctccac 1680
acccgcttca accagacggc caccgccacg ggcaggcttt ccagctccga ccccaacctg 1740
cagaacatcc ccgtgcgcac cctctgggc cagcgcaccc gccgagcctt cgtggccgag 1800
gagggctggg tgctgggtgt cttggactac agccagattg agcttcgggt cctggccac 1860
ctctccgggg acgagaacct gatccgggtc tttcaggagg ggaggacat ccacacccag 1920
accgccagct ggatgttcgg cgtttcccc gaaggggtag accctctgat gcgcggggcg 1980
gccaaagacca tcaacttcgg ggtgctctac ggcattgtcc cccaccgct ctccggggag 2040
ctttccatcc cctacgagga ggcgggtggc ttcatgtgac gctacttcca gagctacccc 2100

Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765

Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His
770 775 780

Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala
785 790 795 800

Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
805 810 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

<210> 5
<211> 831
<212> PRT
<213> thermus flavus

<400> 5
Met Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1 5 10 15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
20 25 30

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
35 40 45

Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Val Val Val Val Val
50 55 60

Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr
65 70 75 80

Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
85 90 95

Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val
100 105 110

Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Arg Ala
115 120 125

Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu
130 135 140

Tyr Gln Leu Leu Ser Glu Arg Ile Ala Ile Leu His Pro Glu Gly Tyr
145 150 155 160

Leu Ile Thr Pro Ala Trp Leu Tyr Glu Lys Tyr Gly Leu Arg Pro Glu
165 170 175

Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile

<211> 836
 <212> PRT
 <213> Thermus thermophilus

<400> 6

Met	Asn	Ser	Glu	Ala	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val
1				5					10					15	
Leu	Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	Phe	Ala	Leu
			20					25					30		
Lys	Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly
		35					40					45			
Phe	Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Tyr	Lys	Ala
	50					55					60				
Val	Phe	Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala
	65				70					75					80
Tyr	Glu	Ala	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro
				85					90					95	
Arg	Gln	Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr
			100					105					110		
Arg	Leu	Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Thr	Leu
		115					120					125			
Ala	Lys	Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala
	130					135					140				
Asp	Arg	Asp	Leu	Tyr	Gln	Leu	Val	Ser	Asp	Arg	Val	Ala	Val	Leu	His
	145				150				155						160
Pro	Glu	Gly	His	Leu	Ile	Thr	Pro	Glu	Trp	Leu	Trp	Glu	Lys	Tyr	Gly
				165					170					175	
Leu	Arg	Pro	Glu	Gln	Trp	Val	Asp	Phe	Arg	Ala	Leu	Val	Gly	Asp	Pro
			180					185					190		
Ser	Asp	Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Leu
		195					200					205			
Lys	Leu	Leu	Lys	Glu	Trp	Gly	Ser	Leu	Glu	Asn	Leu	Leu	Lys	Asn	Leu
	210					215					220				
Asp	Arg	Val	Lys	Pro	Glu	Asn	Val	Arg	Glu	Lys	Ile	Lys	Ala	His	Leu
	225				230					235					240
Glu	Asp	Leu	Arg	Leu	Ser	Leu	Glu	Leu	Ser	Arg	Val	Arg	Thr	Asp	Leu
				245					250					255	
Pro	Leu	Glu	Val	Asp	Leu	Ala	Gln	Gly	Arg	Glu	Pro	Asp	Arg	Glu	Gly
			260					265					270		
Leu	Arg	Ala	Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu
		275					280					285			
Phe	Gly	Leu	Leu	Glu	Ala	Pro	Ala	Pro	Leu	Glu	Glu	Ala	Pro	Trp	Pro
	290					295					300				

Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
 785 790 795 800
 Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
 805 810 815
 Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
 820 825 830
 Ser Ala Lys Gly
 835

<210> 7

<211> 2502

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 7

atgnnggcga tgcttcccct ctttgagccc aaaggccggg tcttcttggt ggacggccac 60
 cacctggcct accgcacctt cttcgccctg aagggcctca ccaccagccg gggcgaaccg 120
 gtgcaggcgg tctacggctt cgccaagagc ctctcaagg cctgaagga ggacggggac 180
 nnggcggtgn tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
 gcctacaagg cgggcccgggc cccacccccg gaggactttc cccggcagct cgccctcatc 300
 aaggagctgg tggacctcct ggggcttgcg cgctcgagg tccccggcta cgaggcggac 360
 gacgtntctg ccacctggc caagaaggcg gaaaaggagg ggtacgaggt gcgcatactc 420

accgccgacc gcgacctcta ccagctcctt tccgaccgca tcgccgtcct ccaccccagag 480
 ggggtacctca tcaccccggc gtggcttttg gagaagtacg gcctgaggcc ggagcagtgg 540
 gtggactacc gggccctggc gggggacccc tccgacaacc tccccggggt caagggcatc 600
 ggggagaaga ccgcccngaa gctcctcnag gagtggggga gcctggaaaa cctcctcaag 660
 aacctggacc ggggtgaagcc cgccttccgg gagaagatcc aggcccatat ggagacctg 720
 angtctcctt gggagctntc ccaggtgcgc accgacctgc ccctggagggt ggacttcgcc 780
 aagnggcggg agcccgaccg ggaggggctt agggcctttc tggagaggct ggagtttggc 840
 agcctcctcc acgagttcgg cctcctggag ggccccaagg ccctggagga ggccccctgg 900
 cccccgccgg aaggggcctt cgtgggcttt gtcctttccc gccccgagcc catgtggggc 960
 gagcttcttg ccctggccgc cgccaggag ggccgggtcc accgggcacc agaccctttt 1020
 angggcctna gggacctnaa ggaggtgcgg ggnctcctcg ccaaggacct ggccgttttg 1080
 gccctgaggg agggcctnga cctcntgccc ggggacgacc ccatgctcct cgctacctc 1140
 ctggacccct ccaacaccac ccccgagggg gtggcccggc gctacggggg ggagtggacg 1200
 gaggangcgg gggagcgggc cctcctntcc gagaggctct tccngaacct nnngcagcgc 1260
 cttgaggggg aggagaggct cctttggctt taccaggagg tggagaagcc cttttcccg 1320
 gtcttgcccc acatggaggc cacgggggtg cggctggacg tggcctacct ccaggccctn 1380
 tccctggagg tggcggagga gatccgccgc ctcgaggagg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagctnggg 1500
 cttcccgcca tcggcaagac ggagaagacn ggcaagcgct ccaccagcgc cgccgtgctg 1560
 gaggccctnc gngaggccca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctcaaga acacctacat ngacccctg ccngncctcg tccacccag gacgggccgc 1680
 ctccacacc gcttcaacca gacggccacg gccacgggca ggcttagtag ctccgacccc 1740
 aacctgcaga acatccccgt ccgcaccccn ctgggcccaga ggatccgccg ggccttcgtg 1800
 gccgaggagg gntgggtgtt ggtggccctg gactatagcc agatagagct ccgggtcctg 1860
 gccacctct ccggggacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
 acccagaccg ccagctggat gttcggcgtc cccccggagg ccgtggaccc cctgatgcgc 1980
 cgggcccga agacatcaa cttcggggtc ctctacggca tgtccgcca ccgcctctcc 2040
 caggagcttg ccatccccta cgaggaggcg gtggccttca ttgagcgcta cttccagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagacc tcttcggccg ccggcgctac gtgcccgaac tcaacgcccg ggtgaagagc 2220
 gtgcgggagg cggcgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280

ctcatgaagc tggccatggt gaagctcttc ccccggtcnc aggaaatggg ggccaggatg 2340
ctcctncagg tccacgacga gctggtcctc gagggcccca aagagcgggc ggaggnggtg 2400
gccgctttgg ccaaggaggt catggagggg gtctatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga tgggggagga ctggctctcc gccaaaggagt ag 2502

<210> 8

<211> 833

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 8

Met	Xaa	Ala	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu	Leu	1	5	10	15
Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	Phe	Ala	Leu	Lys	Gly	20	25	30	
Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe	Ala	35	40	45	
Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Xaa	Val	50	55	60	
Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Glu	Ala	65	70	75	80
Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln	Leu	85	90	95	
Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Xaa	Arg	Leu	Glu	100	105	110	
Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Thr	Leu	Ala	Lys	Lys	115	120	125	
Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Arg	Asp	130	135	140	
Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	Ala	Val	Leu	His	Pro	Glu	Gly	145	150	155	160
Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	Pro	165	170	175	
Glu	Gln	Trp	Val	Asp	Tyr	Arg	Ala	Leu	Xaa	Gly	Asp	Pro	Ser	Asp	Asn	180	185	190	
Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Xaa	Lys	Leu	Leu	195	200	205	
Xaa	Glu	Trp	Gly	Ser	Leu	Glu	Asn	Leu	Leu	Lys	Asn	Leu	Asp	Arg	Val	210	215	220	
Lys	Pro	Xaa	Xaa	Arg	Glu	Lys	Ile	Xaa	Ala	His	Met	Glu	Asp	Leu	Xaa	225	230	235	240

Leu Ser Xaa Xaa Leu Ser Xaa Val Arg Thr Asp Leu Pro Leu Glu Val
 245 250 255
 Asp Phe Ala Xaa Arg Arg Glu Pro Asp Arg Glu Gly Leu Arg Ala Phe
 260 265 270
 Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
 275 280 285
 Glu Xaa Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
 290 295 300
 Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp Ala Glu
 305 310 315 320
 Leu Leu Ala Leu Ala Ala Ala Arg Xaa Gly Arg Val His Arg Ala Xaa
 325 330 335
 Asp Pro Leu Xaa Gly Leu Arg Asp Leu Lys Glu Val Arg Gly Leu Leu
 340 345 350
 Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Leu Asp Leu Xaa
 355 360 365
 Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn
 370 375 380
 Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu
 385 390 395 400
 Asp Ala Gly Glu Arg Ala Leu Leu Ser Glu Arg Leu Phe Xaa Asn Leu
 405 410 415
 Xaa Xaa Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Xaa Glu
 420 425 430
 Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr Gly
 435 440 445
 Val Arg Leu Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Val Ala
 450 455 460
 Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly His
 465 470 475 480
 Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp
 485 490 495
 Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg
 500 505 510
 Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile
 515 520 525
 Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Asn Thr
 530 535 540
 Tyr Ile Asp Pro Leu Pro Xaa Leu Val His Pro Arg Thr Gly Arg Leu
 545 550 555 560
 His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser
 565 570 575

Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln
 580 585 590
 Arg Ile Arg Arg Ala Phe Val Ala Glu Glu Gly Trp Xaa Leu Val Ala
 595 600 605
 Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly
 610 615 620
 Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr
 625 630 635 640
 Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp Pro
 645 650 655
 Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly
 660 665 670
 Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu
 675 680 685
 Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg
 690 695 700
 Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val
 705 710 715 720
 Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
 725 730 735
 Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
 740 745 750
 Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
 755 760 765
 Phe Pro Arg Leu Xaa Glu Met Gly Ala Arg Met Leu Leu Gln Val His
 770 775 780
 Asp Glu Leu Val Leu Glu Ala Pro Lys Xaa Arg Ala Glu Xaa Val Ala
 785 790 795 800
 Ala Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
 805 810 815
 Leu Glu Val Glu Val Gly Xaa Gly Glu Asp Trp Leu Ser Ala Lys Glu
 820 825 830

Xaa

<210> 9
 <211> 53
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 9
 uucgcuuucu ucccuuccuu ucucgccacg uucgccggcu uuccccguca agc

53

<210> 10
 <211> 30
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 10
 acggggaaag ccggcgaacg tggcgagaaa 30

<210> 11
 <211> 26
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 11
 attagaaagg aagggaagaa agcgaa 26

<210> 12
 <211> 30
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 12
 acggggaaag ccggcgaacg tggcgagAAC 30

<210> 13
 <211> 30
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 13
 cttgacgggg aaagccggcg aacgtggcgc 30

<210> 14
 <211> 23
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 14
 agaaaggaag ggaagaaagc gaa 23

<210> 15

<211> 27
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 15
agggagaagg caactggacc gaaggcc

27

<210> 16
<211> 35
<212> DNA
<213> Artificial Sequence

<220>

<221> misc_feature

<222> (1)

<223> This 5' end has a fluorescein label.

<220>

<221> misc_feature

<222> (35)

<223> This 3' end is modified with a dideoxynucleotide.

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 16
ncgaaattaa tacgcttggtg gagaaggagt tcatn

35

<210> 17
<211> 640
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 17
gggagcccag cuaugaacuc cuucuccaca agcgccuucg guccaguugc cuucuccug 60
gggcugcucc ugguguugcc ugcugccuuc ccugccccag uacccccagg agaagauucc 120
aaagauguag ccgccccaca cagacagcca cucaccucu cagaacgaau ugacaaacaa 180
auucgguaca uccucgacgg caucucagcc cugagaaagg agacauguaa caagaguaac 240
augugugaaa gcagcaaaga ggcacuggca gaaaacaacc ugaaccuucc aaagauggcu 300
gaaaaaagau gaugcuucca aucuggauuc aaugaggaga cuugccuggu gaaaaucauc 360
acuggucuuu uggaguuuga gguauaccua gaguaccucc agaacagauu ugagaguagu 420
gaggaacaag ccagagcugu ccagaugagu acaaagucc ugauccaguu ccugcagaaa 480
aaggcaaaga aucuagaugc aaauaccacc ccugacccaa ccacaaaugc cagccugcug 540
acgaagcugc aggcacagaa ccaguggcug caggacauga caacucaucu cauucugcgc 600

agcuuaagg aguuccugca guccagccug agggcucuuc

640

<210> 18
<211> 53
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 18
gctatgaact ccttctccac aagcgccttc ggtccagttg ccttctccct ggg 53

<210> 19
<211> 214
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 19
Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
1 5 10 15
Glu Trp Thr Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu
20 25 30
Phe Ala Asn Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp
35 40 45
Leu Tyr Arg Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met
50 55 60
Glu Ala Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser
65 70 75 80
Leu Glu Val Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg
85 90 95
Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
100 105 110
Val Leu Phe Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys
115 120 125
Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
130 135 140
Ala His Pro Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys
145 150 155 160
Leu Lys Ser Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg
165 170 175
Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
180 185 190
Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
195 200 205

Pro Leu Gly Gln Arg Ile
210

<210> 20
<211> 214
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 20
Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
1 5 10 15
Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
20 25 30
His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
35 40 45
Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
50 55 60
Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
65 70 75 80
Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
85 90 95
Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
100 105 110
Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys
115 120 125
Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
130 135 140
Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
145 150 155 160
Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
165 170 175
Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
180 185 190
Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
195 200 205
Pro Leu Gly Gln Arg Ile
210

<210> 21
<211> 16
<212> DNA
<213> Artificial Sequence

<220>

<221> misc_feature
 <222> (1)
 <223> This 5' end has a fluorescein label.

 <220>
 <221> misc_feature
 <222> (6)
 <223> The residue at this position is a cy3 abasic
 linker group.

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 21
 ncgctntctc gctcgc

16

<210> 22
 <211> 18
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 22
 acggaacgag cgtctttg

18

<210> 23
 <211> 32
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 23
 gcgagcgaga cagcgaaaga cgcucguucc gu

32

<210> 24
 <211> 32
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 24
 gcgagcgaga cagcgaaaga cgctcgttcc gt

32

<210> 25
 <211> 29
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic

 <400> 25

acggaacgag cgtctttcat ctgtcaatc

29

<210> 26
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)
<223> This 5' end is labeled with
tetrachlororfluorescein.

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 26
nttttcaact gccgtga

17

<210> 27
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)
<223> This 5' end is modified with a biotin-streptavidin
complex.

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 27
nucacggcag uuggugcgcc ucggaacgag gcgcacg

37

<210> 28
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 28
tcacggcagt tggcgcgct cggaacgagg cgcacg

36

<210> 29
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (30)
<223> This 3' end is modified with an amine moiety.

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 29
cggaggaagc agttggtgcg cctcgttaan

30

<210> 30
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)
<223> This 5' end is labeled with fluorescein.

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 30
ntccttctca actgcttcct ccg

23

<210> 31
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (28)
<223> This 3' end is modified with a biotin moiety.

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 31
aacgaggcgc acctcaaadc tccctttn

28

<210> 32
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)
<223> This 5' end is labeled with fluorescein.

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 32
nttttcgctg tctcgct

17

<210> 33
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 <400> 33
 acgagcgtct ttg 13

<210> 34
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> misc_feature
 <222> (1)
 <223> This 5' end is labeled with fluorescein.

<220>
 <223> Description of Artificial Sequence: Synthetic
 <400> 34
 nagecagaca gcgaaagacg ctcgtt 26

<210> 35
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 <400> 35
 ucacggcagu uggugcggaa cgcacg 26

<210> 36
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 <400> 36
 tcacggcagt tggcgcggaa cgcacg 26

<210> 37
 <211> 36
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 <400> 37
 cacgaattcg gggatgctgc ccctctttga gcccaa 36

<210> 38
 <211> 34

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 38
gtgagatcta tcactccttg gcggagagcc agtc

34

<210> 39
<211> 2502
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 39
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcgggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgcccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc tccccggggt caagggcatc 600
gggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttcttg ccctggccgc cgccaggggg ggccgggtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgcc ggcgacgacc ccatgctcct cgctacctc 1140
ctggaccctt ccaacaccac ccccagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320

gtcctggccc acatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgggga gatcgcccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc cggggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
 gagggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggcttcatc 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggagc ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc cccgggagg ccgtggacc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttcccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccggggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctccttcagg tccacgacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggagg catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gagtgggga taggggagga ctggctctcc gccaaaggat ga 2502

<210> 40

<211> 833

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 40

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	

Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
		20						25					30		

Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			

Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
405 410 415

Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
420 425 430

Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
435 440 445

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
450 455 460

Ala Gly Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
485 490 495

Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Arg Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu

<210> 41
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 41
cacgaattcc gaggcgatgc ttccgctc 28

<210> 42
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 42
tcgacgtcga ctaacccttg gcggaaagcc 30

<210> 43
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 43
gcatcgctc ggaattcatg gtc 23

<210> 44
<211> 30
<212> DNA
<213> Artificial Sequence

```

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 44
atagccatgg tggagcggcc gctctcccg 30

<210> 45
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 45
aagcgtcgac tcaatcctgc ttgcctcca gcc 33

<210> 46
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 46
aatcgaattc accccacttt ttgacctgga gg 32

<210> 47
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 47
ccgggagagc ggccgctcca c 21

<210> 48
<211> 2508
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 48
atggaattca cccactttt tgacctggag gaacccccca agcgggtgct tctggtggac 60
ggccaccacc tggcctaccg caccttctat gccctgagcc tcaccacctc ccgggggggag 120
ccggtgcaga tggctctacg cttcgcccgg agcctcctca aggccttgaa ggaggacgga 180
caggcggtgg tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
gcctacaagg cgggcccggc cccacccccg gaggacttcc cccgccagct cgccttggtc 300

```

aagcggctgg tggaccttct gggcctggtc cgcctcgagg ccccggggta cgaggcggac 360
gacgtcctgg gcaccctggc caagaaggcc gaaagggagg ggatggaggc gcgcacccctc 420
acgggagacc gggacttctt ccagctcctc tccgagaagg tctcggctct cctgccggac 480
gggaccctgg tcaccccaaa ggacgtccag gagaagtacg gggtgccccc ggagcgctgg 540
gtggacttcc gcgcctccac gggggaccgc tcggacaaca tccccggggt ggcggggata 600
ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
ctccacctct ccttagacct ggcccgcatc cgcaccgacc tccccctgga ggtggacttt 780
aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840
gagttcggaa gcctcctcca cgagttcggc ctctctggag gggagaagcc ccgggaggag 900
gccccctggc ccccgccga aggggccttc gtgggcttcc tcctttcccg caaggagccc 960
atgtgggagg agcttctggc cctggcggcg gcctcggagg gccgggtcca ccgggcaaca 1020
agcccggttg aggccttggc cgacctcaag gagggccggg ggttcctggc caaggacctg 1080
gccgttttgg ccctgcggga ggggggtggc ctggacccca cggacgacct cctcctggtg 1140
gcctacctcc tggaccggc caacaccac cccgaggggg tggcccgcg ctacgggggc 1200
gagttcacgg aggacgcagc ggagagggcc ctctctccg agaggctctt ccagaacctc 1260
tttccccggc tttccgagaa gctcctctgg ctctaccagg aggtggagcg gccctctcc 1320
cgggtcttgg ccacatgga gggccggggg gtgaggctgg acgtccccct tctggaggcc 1380
ctctcctttg agctggagaa ggagatggag cgcctggagg gggaggctct ccgtttggcc 1440
ggccaccct tcaacctcaa ctcccgac cagctggaaa gggcctctt tgacgagctg 1500
ggcctcacc cgggtgggccc gacggagaag acgggcaagc gctccaccgc ccagggggcc 1560
ctggaggccc tccggggggc ccacccatc gtggagctca tcctccagta ccgggagctt 1620
tccaagctca aaagcaccta cctggacccc ctgccccggc tcgtccacc gcggacgggc 1680
cggtccaca cccgttcaa ccagacggc acggccacgg gaaggctttc cagctccgac 1740
cccaacctgc agaacatccc cgtgcgcacc cccttggggc agcgcatccg caaggccttc 1800
gtggccgagg aggggtggct ccttttggcg gcggactact ccagattga gctccgggtc 1860
ctggcccacc tctcggggga cgagaacctg aagcgggtct tccgggaggg gaaggacatc 1920
cataccgaga ccgccgctg gatgttcggc ttagacccc ctctggtgga tccaaagatg 1980
cgccgggagg ccaagacggt caacttcggc gtctctacg ggatgtccgc ccacaggctc 2040
tcccaggagc tcggcataga ctacaaggag gcggaggcct ttattgagcg ctacttccag 2100
agcttcccca aggtgcgggc ctggatagaa aggaccctgg aggagggccg gacgcggggc 2160

tacgtggaga ccctgttcgg caggaggcgc tatgtgcccg acctggcctc ccgggtccgc 2220
tcggtgcggg aggcggcgga gcggatggcc ttcaacatgc ccgtgcaggg caccgccgcc 2280
gacctgatga agatgccat ggtcaagctc tccccaggc taaagcccct gggggccac 2340
ctcctcctcc aagtgcacga cgagctggtc ctggaggtgc ccgaggaccg ggccgaggag 2400
gccaaaggccc tgggtcaagga ggtcatggag aacgcctacc ccctggacgt gccctcgag 2460
gtggaggtgg gcgtgggtcg ggactggctg gaggcgaagc aggattga 2508

<210> 49
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 49
actggaattc ctgcccctct ttgagcccaa g 31

<210> 50
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 50
aacagtcgac ctaggccttg gcggaaagcc 30

<210> 51
<211> 2499
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 51
atggaattcc tgcccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
ctggcctacc gtaccttttt tgccctgaag ggcctcacca ccagccgcgg ggagccggtc 120
caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
gtgatcgtgg tgtttgacgc caaggcccc tccttcgcc accagaccta cgaggcctac 240
aaggcggggc gggctccac ccccaggac tttcccggc agcttgccct tatcaaggag 300
atggtggacc ttttgggcct ggagcgctc gaggtgccgg gctttgaagc ggatgacgtc 360
ctggctaccc tggccaagaa ggcggaaaag gaaggctacg aagtgcgcat cctcaccgcg 420
gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggaggggttac 480

ttgcaggtgc acgacgaact ggtcctcgag gctcccaagg agcaagcgga ggaagtcgcc 2400
caggaggcca agcggaccat ggaggaggtg tggcccctga aggtgccctt ggaggtggaa 2460
gtgggcatcg gggaggactg gctttccgcc aaggcctag 2499

<210> 52
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 52
cgatctcctc ggccacctcc 20

<210> 53
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 53
ggcgggtgccc tggacgggca 20

<210> 54
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 54
ccagctcggtt gtggacctga 20

<210> 55
<211> 2505
<212> DNA
<213> Thermus aquaticus

<220>
<221> CDS
<222> (1)..(2499)

<400> 55
atg aat tcg ggg atg ctg ccc ctc ttt gag ccc aag ggc cgg gtc ctc 48
Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
ctg gtg gac ggc cac cac ctg gcc tac cgc acc ttc cac gcc ctg aag 96
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
ggc ctc acc acc agc cgg ggg gag ccg gtg cag gcg gtc tac ggc ttc 144

Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe		
		35					40					45					
gcc	aag	agc	ctc	ctc	aag	gcc	ctc	aag	gag	gac	ggg	gac	gcg	gtg	atc	192	
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile		
	50					55					60						
gtg	gtc	ttt	gac	gcc	aag	gcc	ccc	tcc	ttc	cgc	cac	gag	gcc	tac	ggg	240	
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly		
	65				70					75					80		
ggg	tac	aag	gcg	ggc	cgg	gcc	ccc	acg	ccg	gag	gac	ttt	ccc	cgg	caa	288	
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln		
				85					90					95			
ctc	gcc	ctc	atc	aag	gag	ctg	gtg	gac	ctc	ctg	ggg	ctg	gcg	cgc	ctc	336	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu		
			100					105					110				
gag	gtc	ccg	ggc	tac	gag	gcg	gac	gac	gtc	ctg	gcc	agc	ctg	gcc	aag	384	
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys		
		115					120					125					
aag	gcg	gaa	aag	gag	ggc	tac	gag	gtc	cgc	atc	ctc	acc	gcc	gac	aaa	432	
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys		
	130					135					140						
gac	ctt	tac	cag	ctc	ctt	tcc	gac	cgc	atc	cac	gtc	ctc	cac	ccc	gag	480	
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu		
	145				150				155						160		
ggg	tac	ctc	atc	acc	ccg	gcc	tgg	ctt	tgg	gaa	aag	tac	ggc	ctg	agg	528	
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg		
				165					170					175			
ccc	gac	cag	tgg	gcc	gac	tac	cgg	gcc	ctg	acc	ggg	gac	gag	tcc	gac	576	
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp		
			180					185					190				
aac	ctt	ccc	ggg	gtc	aag	ggc	atc	ggg	gag	aag	acg	gcg	agg	aag	ctt	624	
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu		
		195					200					205					
ctg	gag	gag	tgg	ggg	agc	ctg	gaa	gcc	ctc	ctc	aag	aac	ctg	gac	cgg	672	
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg		
	210					215					220						
ctg	aag	ccc	gcc	atc	cgg	gag	aag	atc	ctg	gcc	cac	atg	gac	gat	ctg	720	
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu		
	225				230					235					240		
aag	ctc	tcc	tgg	gac	ctg	gcc	aag	gtg	cgc	acc	gac	ctg	ccc	ctg	gag	768	
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu		
				245					250					255			
gtg	gac	ttc	gcc	aaa	agg	cgg	gag	ccc	gac	cgg	gag	agg	ctt	agg	gcc	816	
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala		
			260					265					270				
ttt	ctg	gag	agg	ctt	gag	ttt	ggc	agc	ctc	ctc	cac	gag	ttc	ggc	ctt	864	
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu		
		275					280					285					

ctg gaa agc ccc aag gcc ctg gag gag gcc ccc tgg ccc ccg ccg gaa	912
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu	
290 295 300	
ggg gcc ttc gtg ggc ttt gtg ctt tcc cgc aag gag ccc atg tgg gcc	960
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala	
305 310 315 320	
gat ctt ctg gcc ctg gcc gcc gcc agg ggg ggc cgg gtc cac cgg gcc	1008
Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala	
325 330 335	
ccc gag cct tat aaa gcc ctc agg gac ctg aag gag gcg cgg ggg ctt	1056
Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu	
340 345 350	
ctc gcc aaa gac ctg agc gtt ctg gcc ctg agg gaa ggc ctt ggc ctc	1104
Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu	
355 360 365	
ccg ccc ggc gac gac ccc atg ctc ctc gcc tac ctc ctg gac cct tcc	1152
Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser	
370 375 380	
aac acc acc ccc gag ggg gtg gcc cgg cgc tac ggc ggg gag tgg acg	1200
Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr	
385 390 395 400	
gag gag gcg ggg gag cgg gcc gcc ctt tcc gag agg ctc ttc gcc aac	1248
Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn	
405 410 415	
ctg tgg ggg agg ctt gag ggg gag gag agg ctc ctt tgg ctt tac cgg	1296
Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg	
420 425 430	
gag gtg gag agg ccc ctt tcc gct gtc ctg gcc cac atg gag gcc acg	1344
Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr	
435 440 445	
ggg gtg cgc ctg gac gtg gcc tat ctc agg gcc ttg tcc ctg gag gtg	1392
Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val	
450 455 460	
gcc gag gag atc gcc cgc ctc gag gcc gag gtc ttc cgc ctg gcc ggc	1440
Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly	
465 470 475 480	
cac ccc ttc aac ctc aac tcc cgg gac cag ctg gaa agg gtc ctc ttt	1488
His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe	
485 490 495	
gac gag cta ggg ctt ccc gcc atc ggc aag acg gag aag acc ggc aag	1536
Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys	
500 505 510	
cgc tcc acc agc gcc gcc gtc ctg gag gcc ctc cgc gag gcc cac ccc	1584
Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro	
515 520 525	
atc gtg gag aag atc ctg cag tac cgg gag ctc acc aag ctg aag agc	1632
Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser	

530	535	540	
acc tac att gac ccc ttg ccg gac ctc atc cac ccc agg acg ggc cgc Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg 545 550 555 560			1680
ctc cac acc cgc ttc aac cag acg gcc acg gcc acg ggc agg cta agt Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser 565 570 575			1728
agc tcc gat ccc aac ctc cag aac atc ccc gtc cgc acc ccg ctt ggg Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly 580 585 590			1776
cag agg atc cgc cgg gcc ttc atc gcc gag gag ggg tgg cta ttg gtg Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val 595 600 605			1824
gcc ctg gac tat agc cag ata gag ctc agg gtg ctg gcc cac ctc tcc Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser 610 615 620			1872
ggc gac gag aac ctg atc cgg gtc ttc cag gag ggg cgg gac atc cac Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His 625 630 635 640			1920
acg gag acc gcc agc tgg atg ttc ggc gtc ccc cgg gag gcc gtg gac Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp 645 650 655			1968
ccc ctg atg cgc cgg gcg gcc aag acc atc aac ttc ggg gtc ctc tac Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr 660 665 670			2016
ggc atg tcg gcc cac cgc ctc tcc cag gag cta gcc atc cct tac gag Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu 675 680 685			2064
gag gcc cag gcc ttc att gag cgc tac ttt cag agc ttc ccc aag gtg Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val 690 695 700			2112
cgg gcc tgg att gag aag acc ctg gag gag ggc agg agg cgg ggg tac Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr 705 710 715 720			2160
gtg gag acc ctc ttc ggc cgc cgc cgc tac gtg cca gac cta gag gcc Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala 725 730 735			2208
cgg gtg aag agc gtg cgg gag gcg gcc gag cgc atg gcc ttc aac atg Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met 740 745 750			2256
ccc gtc cag ggc acc gcc gcc gac ctc atg aag ctg gct atg gtg aag Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys 755 760 765			2304
ctc ttc ccc agg ctg gag gaa atg ggg gcc agg atg ctc ctt cag gtc Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val 770 775 780			2352

cac aac gag ctg gtc ctc gag gcc cca aaa gag agg gcg gag gcc gtg	2400
His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val	
785 790 795 800	
gcc cgg ctg gcc aag gag gtc atg gag ggg gtg tat ccc ctg gcc gtg	2448
Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val	
805 810 815	
ccc ctg gag gtg gag gtg ggg ata ggg gag gac tgg ctc tcc gcc aag	2496
Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys	
820 825 830	
gag tgatag	2505
Glu	

<210> 56
 <211> 833
 <212> PRT
 <213> Thermus aquaticus

<400> 56

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu	
1 5 10 15	
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys	
20 25 30	
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe	
35 40 45	
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile	
50 55 60	
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly	
65 70 75 80	
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln	
85 90 95	
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu	
100 105 110	
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys	
115 120 125	
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys	
130 135 140	
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu	
145 150 155 160	
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg	
165 170 175	
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp	
180 185 190	
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu	
195 200 205	
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg	

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu

<210> 57
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 57
caggaggagc tcgttggtgga cctgga

26

<210> 58
<211> 836
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 58
Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220
Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240
Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu

Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
 595 600 605
 Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
 625 630 635 640
 Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
 785 790 795 800
 Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
 805 810 815
 Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
 820 825 830
 Ser Ala Lys Gly
 835

<210> 59

<211> 2511

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 59

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60

```

ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccgggtgc aggcgggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgca cctctaccaa ctctctccg accgcgtcgc cgtcctccac 480
cccgaggggc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcgga gccctctcca cgagttcggc ctctggagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggagcg gccgggtgca ccgggcagca 1020
gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgcccg gggacgacct catgctctc 1140
gcctacctcc tggaccctc caacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
gagtggacgg aggacgcgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggcca catggaggcc accggggtac ggcgggacgt ggctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
acgggccgcc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgaccca acctgcagaa catccccgtc cgcacccct tgggcccagag gatccgccg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
cgcgtcctcg ccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920

```

gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccgaggc cgtggacccc 1980
 ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccac 2040
 aggtctctccc aggagcttgc catccccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tccccaaagt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct ctctcggaaga aggcgctacg tgcccgcact caacgcccgg 2220
 gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gccgcgatgc tcctccaggt ccacaacgag ctctctctgg agggccccca agcgcggggc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtta g 2511

<210> 60
 <211> 50
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 60
 tctagaggat ctatcagtgg tgggtggtgg ggtgctcctt ggcggagagc 50

<210> 61
 <211> 58
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 61
 tgcctgcagg tcgacgctag ctagtgggtg tgggtggtgg gacccttggc ggaaagcc 58

<210> 62
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 62
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cgggtctacg ctctcgccaag agcctcctca aggcctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcttacggg 240

ggggtacaagg cgggcccgggc cccacgccc gaggaactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgccc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 cccccgccgg aaggggcctt cgtgggcctt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggg ggccgggtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctccgccc ggcgacgacc ccatgctcct cgctacctc 1140
 ctggacctt ccaacaccac ccccaggggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtcttgcccc acatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgcccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac ggagaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccttgg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggacc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcgccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100

ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg 2340
 ctcccttcagg tccacaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 63
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 63
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 64
 <211> 2526
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 64

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcgg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccaggg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgagggtccc cggctacgag 360
gcggaacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgagggtgcgc 420
atcctcaccg ccgaccgcga cctctaccaa ctctgtctccg accgcgtcgc cgtcctccac 480
cccagggggc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagagggtg 840
gagttcggca gcctcctcca cgagttcggc ctctggagg cccccgccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
atgtgggagg agcttaaagc cctggccgcc tgcagggacg gccgggtgca ccgggcagca 1020
gacccttggt cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgtccc gggacgacct catgtcctc 1140
gcctacctcc tggacctctc caacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgagg ggaaaagccc 1320
ctctcccggg tcctggccca catggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggccttacg ggaggccac ccatcgtgg agaagatcct ccagcaccgg 1620
gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
acgggccgcc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740

tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg 1800
 gccttctgtg ccgaggcggg ttgggcgttg gtggcccttg actatagcca gatagagctc 1860
 cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctttcca ggaggggaag 1920
 gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcggccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac 2040
 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tccccaaagt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccacac caacgcccgg 2220
 gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
 gcccgcacgc tctccaggt ccacaacgag ctctcctctg agggccccca agcgcgggcc 2400
 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
 ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
 caccac 2526

<210> 65

<211> 842

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 65

Met	Asn	Ser	Glu	Ala	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val
1				5					10					15	
Leu	Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	Phe	Ala	Leu
			20					25					30		
Lys	Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly
		35					40					45			
Phe	Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Tyr	Lys	Ala
		50				55					60				
Val	Phe	Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala
	65				70					75				80	
Tyr	Glu	Ala	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro
			85						90					95	
Arg	Gln	Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr
		100						105					110		
Arg	Leu	Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Thr	Leu
		115					120						125		

Leu	Glu	Leu	Ala	Glu	Glu	Ile	Arg	Arg	Leu	Glu	Glu	Glu	Val	Phe	Arg	
465					470					475					480	
Leu	Ala	Gly	His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	
				485					490						495	
Val	Leu	Phe	Asp	Glu	Leu	Arg	Leu	Pro	Ala	Leu	Gly	Lys	Thr	Gln	Lys	
			500					505					510			
Thr	Gly	Lys	Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	
		515					520					525				
Ala	His	Pro	Ile	Val	Glu	Lys	Ile	Leu	Gln	His	Arg	Glu	Leu	Thr	Lys	
	530					535					540					
Leu	Lys	Asn	Thr	Tyr	Val	Asp	Pro	Leu	Pro	Ser	Leu	Val	His	Pro	Arg	
545					550					555					560	
Thr	Gly	Arg	Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	
				565					570						575	
Arg	Leu	Ser	Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	
			580					585					590			
Pro	Leu	Gly	Gln	Arg	Ile	Arg	Arg	Ala	Phe	Val	Ala	Glu	Ala	Gly	Trp	
		595					600					605				
Ala	Leu	Val	Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	
	610					615					620					
His	Leu	Ser	Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Lys	
625					630					635					640	
Asp	Ile	His	Thr	Gln	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Pro	Glu	
				645					650					655		
Ala	Val	Asp	Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Val	Asn	Phe	Gly	
			660					665					670			
Val	Leu	Tyr	Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	
		675					680					685				
Pro	Tyr	Glu	Glu	Ala	Val	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	
	690					695					700					
Pro	Lys	Val	Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Lys	
705					710					715					720	
Arg	Gly	Tyr	Val	Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	
				725					730					735		
Leu	Asn	Ala	Arg	Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	
			740					745					750			
Phe	Asn	Met	Pro	Val	Gln	Gly	Thr	Ala	Ala	Asp	Leu	Met	Lys	Leu	Ala	
		755					760					765				
Met	Val	Lys	Leu	Phe	Pro	Arg	Leu	Arg	Glu	Met	Gly	Ala	Arg	Met	Leu	
	770					775					780					
Leu	Gln	Val	His	Asn	Glu	Leu	Leu	Leu	Glu	Ala	Pro	Gln	Ala	Arg	Ala	
785					790					795					800	

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 66
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 66
gccgccaggg gcggccgcgt ccaccggggc 30

<210> 67
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 67
gcctgcaggg gcggccgcgt gcaccggggc a 31

<210> 68
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 68
ctcctggacc cttcgaacac cacc 26

<210> 69
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 69
gtcctggccc atatggaggc cac 23

<210> 70
<211> 2526
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 70

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttccg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgcga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gacccttg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttg cctcgaggga ggggctagac ctggtgccc gggacgacct catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctccccgg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgc tcgaggagga ggtcttccgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggccttacg ggaggccac cccatcgagg agaagatcct ccagcaccgg 1620
gagctacca agctcaagaa cacctacgtg gaccctcc caagcctcgt ccacccgagg 1680
acgggccgcc tocacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740

```
tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggcccttg actatagcca gatagagctc 1860
cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccgagggc cgtggacccc 1980
ctgatgcgcc gggcgggcaa gacgggtgaac ttcggcgtcc tctacggcat gtccgcccac 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaggt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaacctt cttcggaaga aggcgttacg tgcccgcact caacgccccg 2220
gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcacgc tcctccaggt ccacaacgag ctctccttgg agggccccca agcgcggggc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggta ccaaccaccac 2520
caccac 2526
```

```
<210> 71
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
```

```
<400> 71
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctcctct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacg cttcgccaag agcctcctca aggcctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc ccccacgccg gaggacttcc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
```


aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgcaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccgcc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccttttcc gagaggctct tcgccaacct gtgggggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac ggagaagacc ggcaagcgt ccaccagcgc cgcctcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccttgg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcctc 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggacct cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagacct tcttcggccc ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcggac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
gcccggtctg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 72

<211> 842
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 72

```

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
 1           5           10           15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
          20           25           30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
          35           40           45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
          50           55           60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
          65           70           75           80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
          85           90           95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
          100          105          110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
          115          120          125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
          130          135          140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
          145          150          155          160

Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
          165          170          175

Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
          180          185          190

Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
          195          200          205

Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
          210          215          220

Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
          225          230          235          240

Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
          245          250          255

Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
          260          265          270

Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
          275          280          285

```

Phe	Gly	Leu	Leu	Glu	Ala	Pro	Ala	Pro	Leu	Glu	Glu	Ala	Pro	Trp	Pro	290	295	300
Pro	Pro	Glu	Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Pro	Glu	Pro	305	310	315
Met	Trp	Ala	Glu	Leu	Lys	Ala	Leu	Ala	Ala	Cys	Arg	Gly	Gly	Arg	Val	325	330	335
His	Arg	Ala	Ala	Asp	Pro	Leu	Ala	Gly	Leu	Lys	Asp	Leu	Lys	Glu	Val	340	345	350
Arg	Gly	Leu	Leu	Ala	Lys	Asp	Leu	Ala	Val	Leu	Ala	Ser	Arg	Glu	Gly	355	360	365
Leu	Asp	Leu	Val	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	370	375	380
Asp	Pro	Ser	Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	385	390	395
Glu	Trp	Thr	Glu	Asp	Ala	Ala	His	Arg	Ala	Leu	Leu	Ser	Glu	Arg	Leu	405	410	415
His	Arg	Asn	Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Lys	Leu	Leu	Trp	420	425	430
Leu	Tyr	His	Glu	Val	Glu	Lys	Pro	Leu	Ser	Arg	Val	Leu	Ala	His	Met	435	440	445
Glu	Ala	Thr	Gly	Val	Arg	Arg	Asp	Val	Ala	Tyr	Leu	Gln	Ala	Leu	Ser	450	455	460
Leu	Glu	Leu	Ala	Glu	Glu	Ile	Arg	Arg	Leu	Glu	Glu	Glu	Val	Phe	Arg	465	470	475
Leu	Ala	Gly	His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	485	490	495
Val	Leu	Phe	Asp	Glu	Leu	Arg	Leu	Pro	Ala	Leu	Gly	Lys	Thr	Gln	Lys	500	505	510
Thr	Gly	Lys	Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	515	520	525
Ala	His	Pro	Ile	Val	Glu	Lys	Ile	Leu	Gln	His	Arg	Glu	Leu	Thr	Lys	530	535	540
Leu	Lys	Asn	Thr	Tyr	Val	Asp	Pro	Leu	Pro	Ser	Leu	Val	His	Pro	Arg	545	550	555
Thr	Gly	Arg	Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	565	570	575
Arg	Leu	Ser	Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	580	585	590
Pro	Leu	Gly	Gln	Arg	Ile	Arg	Arg	Ala	Phe	Val	Ala	Glu	Ala	Gly	Trp	595	600	605
Ala	Leu	Val	Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	610	615	620

Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	385	390	395	400
Glu	Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	405	410	415	
Leu	Trp	Gly	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	420	425	430	
Glu	Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	435	440	445	
Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	450	455	460	
Ala	Glu	Glu	Ile	Ala	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly	465	470	475	480
His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe	485	490	495	
Asp	Glu	Leu	Gly	Leu	Pro	Ala	Ile	Gly	Lys	Thr	Glu	Lys	Thr	Gly	Lys	500	505	510	
Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro	515	520	525	
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser	530	535	540	
Thr	Tyr	Ile	Asp	Pro	Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg	545	550	555	560
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	565	570	575	
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly	580	585	590	
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val	595	600	605	
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser	610	615	620	
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His	625	630	635	640
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp	645	650	655	
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr	660	665	670	
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu	675	680	685	
Glu	Ala	Gln	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val	690	695	700	
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr	705	710	715	720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 74

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 74

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120

ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180

gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240

gggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300

aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360

gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420

accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480

gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540

gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600

ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660

aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720

aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780

aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 75

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala

305		310		315		320
Asp Leu Leu Ala	Leu Ala Ala Ala Arg	Gly Gly Arg Val His Arg Ala				
	325		330		335	
Ala Asp Pro Leu	Ala Gly Leu Lys Asp	Leu Lys Glu Val Arg Gly Leu				
	340		345		350	
Leu Ala Lys Asp	Leu Ala Val Leu Ala Ser Arg	Glu Gly Leu Asp Leu				
	355		360		365	
Val Pro Gly Asp	Asp Pro Met Leu Leu Ala Tyr	Leu Leu Asp Pro Ser				
	370		375		380	
Asn Thr Thr Pro	Glu Gly Val Ala Arg Arg	Tyr Gly Gly Glu Trp Thr				
	385		390		395	400
Glu Asp Ala Ala	His Arg Ala Leu Leu Ser	Glu Arg Leu His Arg Asn				
	405		410		415	
Leu Leu Lys Arg	Leu Glu Gly Glu Glu Lys	Leu Leu Trp Leu Tyr His				
	420		425		430	
Glu Val Glu Lys	Pro Leu Ser Arg Val Leu Ala His	Met Glu Ala Thr				
	435		440		445	
Gly Val Arg Arg	Asp Val Ala Tyr Leu Gln Ala	Leu Ser Leu Glu Leu				
	450		455		460	
Ala Glu Glu Ile	Arg Arg Leu Glu Glu Glu Val	Phe Arg Leu Ala Gly				
	465		470		475	480
His Pro Phe Asn	Leu Asn Ser Arg Asp Gln	Leu Glu Arg Val Leu Phe				
	485		490		495	
Asp Glu Leu Arg	Leu Pro Ala Leu Gly Lys	Thr Gln Lys Thr Gly Lys				
	500		505		510	
Arg Ser Thr Ser	Ala Ala Val Leu Glu Ala Leu Arg	Glu Ala His Pro				
	515		520		525	
Ile Val Glu Lys	Ile Leu Gln His Arg Glu Leu Thr	Lys Leu Lys Asn				
	530		535		540	
Thr Tyr Val Asp	Pro Leu Pro Ser Leu Val His	Pro Arg Thr Gly Arg				
	545		550		555	560
Leu His Thr Arg	Phe Asn Gln Thr Ala Thr Ala Thr	Gly Arg Leu Ser				
	565		570		575	
Ser Ser Asp Pro	Asn Leu Gln Asn Ile Pro Val Arg	Thr Pro Leu Gly				
	580		585		590	
Gln Arg Ile Arg	Arg Ala Phe Val Ala Glu Ala Gly	Trp Ala Leu Val				
	595		600		605	
Ala Leu Asp Tyr	Ser Gln Ile Glu Leu Arg Val Leu	Ala His Leu Ser				
	610		615		620	
Gly Asp Glu Asn	Leu Ile Arg Val Phe Gln Glu Gly	Lys Asp Ile His				
	625		630		635	640

Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
785 790 795 800

Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly His His His His His His
835

<210> 76

<211> 2526

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 76

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60

ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120

gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180

gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcgc ccacgaggcc 240

tacgaggcct acaaggcggg gagggccccg acccccagag acttccccgc gcagctcgcc 300

ctcatcaagg agctggtgga cctcctgggg ttaccgcgc tcgaggtccc cggtacgag 360

gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420

atcctcaccg ccgaccgga cctctacca ctcgtctccg accggtcgc cgtcctccac 480
cccaggggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggaggtg 780
gacctcgccc aggggcggga gcccagaccg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctctggagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccggtcca ccgggcccc 1020
gagccttata aagccctcag ggacctgaag gaggcgcggg ggcttctcgc caaagacctg 1080
agcgttctgg ccctgaggga aggccttggc ctcccgcccc gcgacgacct catgctctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcccgggc ctacggcggg 1200
gagtggacgg aggaggcggg ggagcgggcc gccctttccg agaggctctt cgccaacctg 1260
tgggggaggc ttgaggggga ggagaggctc ctttggcttt accgggaggt ggagaggccc 1320
ctttccgctg tcctggccca tatggaggcc acgggggtgc gcctggacgt ggcctatctc 1380
agggccttgt ccctggaggt ggccgaggag atcgcccgcc tcgaggccga ggtcttccgc 1440
ctggccggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt cctctttgac 1500
gagctagggc tttccgccat cggcaagacg gagaagaccg gcaagcgctc caccagcgcc 1560
gccgtcctgg aggcctcctc cgaggccac cccatcgtgg agaagatcct gcagtaccg 1620
gagctacca agctgaagag caoctacatt gaccttgc cgacctcat ccacccag 1680
acgggccgcc tccacaccg cttcaaccag acggccacgg ccacgggcag gctaagtagc 1740
tccgatcca acctccagaa catccccgtc cgcacccgc ttgggcagag gatccgccc 1800
gccttcatcg ccgaggagg gtggctattg gtggccctgg actatagcca gatagagctc 1860
aggtgctgg cccacctctc cggcgacgag aacctgatcc gggcttcca ggagggcg 1920
gacatccaca cggagaccgc cagctggatg ttcggcgctc cccgggaggc cgtggacccc 1980
ctgatgcgcc gggcgccaa gaccatcaac ttcggggtcc tctacggcat gtcggccac 2040
cgctctccc aggagctagc catcccttac gaggaggccc aggccttcat tgagcgctac 2100
tttcagagct tccccagggt gcgggcctgg attgagaaga ccctggagga gggcaggagg 2160
cgggggtacg tggagacct cttcgccgc cgccgctacg tgccagacct agaggcccg 2220
gtgaagagcg tgcgggaggc ggccgagcgc atggccttca acatgcccg ccagggcacc 2280

```

gccgccgacc tcatgaagct ggctatggtg aagctcttcc ccaggctgga ggaaatgggg 2340
gccaggatgc tccttcaggt ccacaacgag ctggtcctcg aggcccaaaa agagagggcg 2400
gaggccgtgg cccggctggc caaggaggtc atggaggggg tgtatcccct ggccgtgccc 2460
ctggaggtgg aggtggggat aggggaggac tggtctctcg ccaaggagca ccaccaccac 2520
caccac 2526

```

```

<210> 77
<211> 842
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic

```

```

<400> 77
Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
  1             5             10             15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
      20             25             30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
      35             40             45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
      50             55             60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
      65             70             75             80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
      85             90             95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
      100            105            110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
      115            120            125
Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
      130            135            140
Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
      145            150            155            160
Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
      165            170            175
Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
      180            185            190
Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
      195            200            205
Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
      210            215            220

```


Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575
 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp
 595 600 605
 Leu Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg
 625 630 635 640
 Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala
 785 790 795 800
 Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro
 805 810 815
 Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu
 820 825 830
 Ser Ala Lys Glu His His His His His His
 835 840

<210> 78

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

gccacctct ccggcgacga gaacctgac cggtctctcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccacccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 79
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 79
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160

Asp	Glu	Leu	Arg	Leu	Pro	Ala	Leu	Gly	Lys	Thr	Gln	Lys	Thr	Gly	Lys		
			500					505					510				
Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro		
		515					520					525					
Ile	Val	Glu	Lys	Ile	Leu	Gln	His	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Asn		
	530					535					540						
Thr	Tyr	Val	Asp	Pro	Leu	Pro	Ser	Leu	Val	His	Pro	Arg	Thr	Gly	Arg		
545					550					555					560		
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser		
				565					570					575			
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly		
			580					585					590				
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val		
		595					600					605					
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser		
	610					615					620						
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His		
625					630					635					640		
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp		
				645					650					655			
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr		
			660					665					670				
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu		
		675					680					685					
Glu	Ala	Gln	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val		
	690					695					700						
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr		
705					710					715					720		
Val	Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	Leu	Glu	Ala		
				725					730					735			
Arg	Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	Phe	Asn	Met		
			740					745					750				
Pro	Val	Gln	Gly	Thr	Ala	Ala	Asp	Leu	Met	Lys	Leu	Ala	Met	Val	Lys		
		755					760					765					
Leu	Phe	Pro	Arg	Leu	Glu	Glu	Met	Gly	Ala	Arg	Met	Leu	Leu	Gln	Val		
	770					775					780						
His	Asn	Glu	Leu	Val	Leu	Glu	Ala	Pro	Lys	Glu	Arg	Ala	Glu	Ala	Val		
785					790					795					800		
Ala	Arg	Leu	Ala	Lys	Glu	Val	Met	Glu	Gly	Val	Tyr	Pro	Leu	Ala	Val		
				805					810					815			
Pro	Leu	Glu	Val	Glu	Val	Gly	Ile	Gly	Glu	Asp	Trp	Leu	Ser	Ala	Lys		
			820					825					830				

Glu His His His His His His
835

<210> 80
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 80
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggcttacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc ctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg cttcttgaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cttcccggcc ggcgacgacc ccattgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
cttgaggggg aggagaggct cttttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500

```

cttccccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gagggccctcc gcgaggccca ccccatcggtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacggggccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcgtg 1800
gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc 1860
gccacctct ccgggggacga aaacctgatc aggggtcttc aggaggggaa ggacatccac 1920
accagaccg caagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcgccca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc 2040
caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
gtggaaaccc tcttcggaag aaggcgctac gtgcccgcacc tcaacgcccg ggtgaagagc 2220
gtcagggagg ccgcggagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tcgccatggt gaagctcttc cccgcctcc gggagatggg ggcccgcacg 2340
ctcctccagg tccacaacga gtcctcctg gagggccccc aagcgcgggc cgaggaggtg 2400
gcggcctttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggaggtg 2460
gaggtgggga tgggggagga ctggccttcc gccaaagggtc accaccacca ccaccac 2517

```

<210> 81

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 81

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1             5             10             15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
      20             25             30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
      35             40             45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
      50             55             60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
      65             70             75             80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
      85             90             95

```

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430

Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
 785 790 795 800
 Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Gly His His His His His His
 835

<210> 82
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 82
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cggcttacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 gggtagctca tcaccccgcc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caagggtgcg accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg cttcttgaa agccccaagg ccctggagga ggccccctgg 900
 cccccgccg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgcc ggggacgacc ccatgctcct cgctacctc 1140

Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe	
		35					40					45				
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile	
	50					55					60					
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly	
	65				70					75					80	
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln	
				85					90					95		
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu	
			100					105					110			
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys	
		115					120					125				
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys	
	130					135					140					
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu	
	145				150					155					160	
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	
			165						170					175		
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	
		180						185					190			
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu	
		195					200					205				
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg	
	210					215					220					
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu	
	225				230					235					240	
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	
			245					250						255		
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	
			260					265					270			
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	
		275					280					285				
Leu	Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	
	290					295					300					
Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	
	305				310					315					320	
Asp	Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	
			325						330					335		
Ala	Asp	Pro	Leu	Ala	Gly	Leu	Lys	Asp	Leu	Lys	Glu	Val	Arg	Gly	Leu	
			340					345					350			
Leu	Ala	Lys	Asp	Leu	Ala	Val	Leu	Ala	Ser	Arg	Glu	Gly	Leu	Asp	Leu	
		355					360					365				

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 84

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 84

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccgggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggaggt ggacttcgcc 780

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 85

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 86

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 86

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120

ccggtgcagg cggcttacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180

gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240

gggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300

aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360

gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatacctc 420

accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
 gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
 tccttgagc ttgcggagga gatccgccc ctcgaggagg aggtcttccg cttggcgggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
 cttcccgcc tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
 gaggccctac gggaggccca cccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
 aagctcaaga acacctacgt ggacccctc ccaagcctcg tccaccgag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
 aacctgcaga acatccccgt ccgcacccc ttgggccaaga ggatccgccc ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 tcccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgcccagc 2280

ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggcogtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 87

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 87

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
			35				40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
65					70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
145				150						155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
			165						170					175	
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp
			180					185					190		
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu
		195					200					205			
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg
	210					215					220				
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu
225				230						235					240

[illegible]

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 88

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 88

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccggg cccacgcgcg gaggactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
 gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc ctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggcgcgtgc accgggcagc agaccccttg 1020
 gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgcc ggggacgacc ccattgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
 cttgaggggg aggagaggct cttttggctt taccgggagg tggagaggcc cttttccgct 1320
 gtctggccc atatggaggc caggggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac ggagaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccttgg ccggacctca tccaccccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcatc 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct ccggcgacga gaacctgate cgggtcttcc aggagggggc ggacatccac 1920

acggagaccg ccagctggat gttcggcgctc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcgccca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgcccac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 89
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 89
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
325 330 335

Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
340 345 350

Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
355 360 365

Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
370 375 380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
385 390 395 400

Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
405 410 415

Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
420 425 430

Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
435 440 445

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
450 455 460

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
485 490 495

Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 90
<211> 2526
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 90
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaagccc cctccttccg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccagag acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgccg tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgcga cctctaccaa ctcgctctcc accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcgga gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctcgtgccc gggacgacct catgctctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcggg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggcca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaaggg gctctttgac 1500
gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560

gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctacca agtcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg 1680
acgggcccgc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgacccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgcccg 1800
gccttcatcg ccgaggagg gtggctattg gtggccctgg actatagcca gatagagctc 1860
aggggtgctgg cccacctctc cggcgacgag aacctgatcc ggggtctcca ggaggggagg 1920
gacatccaca cggagaccgc cagctggatg ttcggcgctc cccgggaggc cgtggacccc 1980
ctgatgcgcc gggcgccaa gaccatcaac ttcgggggtcc tctacggcat gtcggccccc 2040
cgcctctccc aggagctagc catcccttac gaggaggccc aggccttcat tgagcgctac 2100
tttcagagct tccccagggt gcgggcctgg attgagaaga ccctggagga gggcaggagg 2160
cgggggtacg tggagaccct ctctggccgc cgccgctacg tgccagacct agaggcccgg 2220
gtgaagagcg tgcgggaggc ggccgagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct ggctatggtg aagctcttcc ccaggctgga ggaaatgggg 2340
gccaggatgc tccttcaggt ccacaacgag ctggtcctcg agggcccaaa agagagggcg 2400
gaggccgtgg cccggctggc caaggaggtc atggaggggg tgtatcccct ggccgtgccc 2460
ctggaggtgg aggtggggat aggggaggac tggctctccg ccaaggagca ccaccaccac 2520
caccac 2526

<210> 91
<211> 842
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 91
Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95

Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala
785 790 795 800

Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Glu His His His His His His
835 840

<210> 92
<211> 2526
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 92
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg cctctctcgt ggtctttgac gccaaaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccgagg acttccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgcga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480
cccgagggcc acctcatcac cccggagtggt ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagagggtg 840
gagttcgga gctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcagggcg gccgcgtcca ccgggcccc 1020
gagccttata aagccctcag ggacctgaag gaggcgcggg ggcttctcgc caaagacctg 1080
agcgttctgg ccctgaggga aggccttggc ctcccggccc gcgacgacct catgctcctc 1140

gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcccggcg ctacggcggg 1200
gagtggacgg aggaggcggg ggagcggggc gccctttccg agaggctctt cgccaacctg 1260
tgggggaggg ttgaggggga ggagaggtct ctttggcttt accgggaggt ggagaggccc 1320
ctttccgctg tcctggccca tatggaggcc acgggggtgc gcctggacgt ggcctatctc 1380
agggccttgt ccctggaggt ggccgaggag atcgcccgcg tcgaggccga ggtcttccgc 1440
ctggccgggc accccttcaa cctcaactcc cgggaccagc tggaaagggt cctctttgac 1500
gagctagggc ttcccgccat cggcaagacg gagaagaccg gcaagcgctc caccagcgcc 1560
gccgtcctgg aggccctccg cgaggccccc cccatcgtag agaagatcct gcagtaccgg 1620
gagctcacca agctgaagag cacctacatt gaccccttgc cggacctcat ccaccccagg 1680
acgggcccgc tccacaccgc cttcaaccag acggccacgg ccacgggcag gctaagtagc 1740
tccgatccca acctccagaa catccccgtc cgcaccccgc ttgggcagag gatccgcccg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
cgcgctctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatccaca cccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc 1980
ctgatgcgcc gggcgggcaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac 2040
aggetctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaggt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccacact caacgcccgg 2220
gtgaagagcg tcagggaggg cgcgagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcacgc tcctccaggt ccacaacgag ctctccttgg agggccccc aagcgggggc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
caccac 2526

<210> 93

<211> 842

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 93

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 94

<211> 2499

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 94

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120

ccggtgcagg cggctctacg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180

gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240

gggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgcctcatc 300

aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360

gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420

accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480

gggtacctca tcaccccgcc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540

gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600

ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660

aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgccgc gctacggggg ggagtggacg 1200
 gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cttttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
 tccttgagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaagg tgctctttga cgagcttagg 1500
 cttccgcct tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
 gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
 aagctcaaga acacctacgt ggacccccct ccaagcctcg tccacccgag gacgggccgc 1680
 ctccacacc gttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
 aacctgcaga acatccccgt ccgcaccccc ttggggccaga ggatccgccg ggccttcgtg 1800
 gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc 1860
 gccacctct ccggggacga aaacctgatc aggtcttcc aggggggaa ggacatccac 1920
 acccagaccg caagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cgggcggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc 2040
 caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
 ttcccgaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
 gtggaaaccc tcttcggaag aaggcgctac gtgcccgacc tcaacgcccg ggtgaagagc 2220
 gtcaggaggg ccgcggagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tcgccatggt gaagctcttc cccgcctcc gggagatggg ggcccgcagc 2340
 ctctccagg tccacgacga gctcctcctg gaggcccccc aagcgccggc cgaggaggtg 2400
 gcggcttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggaggtg 2460
 gaggtgggga tgggggagga ctggctttcc gccaaagggt 2499

<210> 95
 <211> 833
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 95

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
65					70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
145					150					155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
			165					170						175	
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp
			180					185					190		
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu
	195						200					205			
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg
	210					215					220				
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu
225					230					235					240
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu
			245						250					255	
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala
			260					265					270		
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu
		275					280					285			

Leu 290	Ser	Pro	Lys	Ala	Leu 295	Glu	Glu	Ala	Pro	Trp 300	Pro	Pro	Pro	Glu	
Gly 305	Ala	Phe	Val	Gly	Phe 310	Val	Leu	Ser	Arg	Lys 315	Glu	Pro	Met	Trp	Ala 320
Asp	Leu	Leu	Ala	Leu 325	Ala	Ala	Ala	Arg	Gly 330	Gly	Arg	Val	His	Arg 335	Ala
Ala	Asp	Pro	Leu 340	Ala	Gly	Leu	Lys	Asp 345	Leu	Lys	Glu	Val	Arg 350	Gly	Leu
Leu	Ala	Lys 355	Asp	Leu	Ala	Val	Leu 360	Ala	Ser	Arg	Glu	Gly 365	Leu	Asp	Leu
Val	Pro 370	Gly	Asp	Asp	Pro	Met 375	Leu	Leu	Ala	Tyr	Leu 380	Leu	Asp	Pro	Ser
Asn 385	Thr	Thr	Pro	Glu	Gly 390	Val	Ala	Arg	Arg	Tyr 395	Gly	Gly	Glu	Trp	Thr 400
Glu	Asp	Ala	Ala	His 405	Arg	Ala	Leu	Leu	Ser 410	Glu	Arg	Leu	His	Arg 415	Asn
Leu	Leu	Lys 420	Arg	Leu	Glu	Gly	Glu	Glu 425	Lys	Leu	Leu	Trp	Leu 430	Tyr	His
Glu	Val	Glu 435	Lys	Pro	Leu	Ser	Arg 440	Val	Leu	Ala	His	Met 445	Glu	Ala	Thr
Gly 450	Val	Arg	Arg	Asp	Val	Ala 455	Tyr	Leu	Gln	Ala	Leu 460	Ser	Leu	Glu	Leu
Ala 465	Glu	Glu	Ile	Arg	Arg 470	Leu	Glu	Glu	Glu	Val 475	Phe	Arg	Leu	Ala	Gly 480
His	Pro	Phe	Asn	Leu 485	Asn	Ser	Arg	Asp	Gln 490	Leu	Glu	Arg	Val	Leu 495	Phe
Asp	Glu	Leu	Arg 500	Leu	Pro	Ala	Leu	Gly 505	Lys	Thr	Gln	Lys	Thr 510	Gly	Lys
Arg	Ser	Thr 515	Ser	Ala	Ala	Val	Leu 520	Glu	Ala	Leu	Arg	Glu 525	Ala	His	Pro
Ile 530	Val	Glu	Lys	Ile	Leu	Gln 535	His	Arg	Glu	Leu	Thr 540	Lys	Leu	Lys	Asn
Thr 545	Tyr	Val	Asp	Pro	Leu 550	Pro	Ser	Leu	Val	His 555	Pro	Arg	Thr	Gly	Arg 560
Leu	His	Thr	Arg	Phe 565	Asn	Gln	Thr	Ala	Thr 570	Ala	Thr	Gly	Arg	Leu 575	Ser
Ser	Ser	Asp	Pro 580	Asn	Leu	Gln	Asn	Ile 585	Pro	Val	Arg	Thr	Pro 590	Leu	Gly
Gln	Arg	Ile 595	Arg	Arg	Ala	Phe	Val 600	Ala	Glu	Ala	Gly	Trp 605	Ala	Leu	Val
Ala 610	Leu	Asp	Tyr	Ser	Gln 615	Ile	Glu	Leu	Arg	Val	Leu 620	Ala	His	Leu	Ser

Figure 1 shows the results of the first two steps of the analysis. The first step, a principal component analysis, revealed that the 10 items loaded on two factors. The first factor, labeled 'Attitudes', included items 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. The second factor, labeled 'Behaviors', included items 11, 12, 13, 14, 15, 16, 17, 18, 19, and 20. The second step, a confirmatory factor analysis, confirmed the two-factor structure. The fit indices for the two-factor model were excellent, indicating that the two-factor model was a good fit for the data. The factor loadings for the two factors were also high, indicating that the items were strongly related to their respective factors.

<220>
<223> Description of Artificial Sequence: Synthetic

```
<400> 96
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgcceaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgcc cccacgccg gaggaacttt cccqgcaact cggcctcctc 300
```

aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcacctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg 1020
gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgcc ggggacgacc ccattgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccg 1320
gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
tccttgagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
cttccgcct tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
aagctcaaga acacctacgt ggacccccctc ccaagcctcg tccaccgag gacgggccgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
aacctgcaga acatccccgt ccgcaccccc ttggggcaga ggatccgccg ggcttccatc 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgac cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgct ccccgggagg ccgtggacc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160

gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccc tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctccctcagg tccacgacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gcccaaggag 2499

<210> 97
 <211> 833
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 97
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg

	210					215					220					
Leu 225	Lys	Pro	Ala	Ile	Arg 230	Glu	Lys	Ile	Leu	Ala 235	His	Met	Asp	Asp	Leu 240	
Lys	Leu	Ser	Trp	Asp 245	Leu	Ala	Lys	Val	Arg 250	Thr	Asp	Leu	Pro	Leu 255	Glu	
Val	Asp	Phe	Ala 260	Lys	Arg	Arg	Glu	Pro 265	Asp	Arg	Glu	Arg	Leu 270	Arg	Ala	
Phe	Leu	Glu 275	Arg	Leu	Glu	Phe	Gly 280	Ser	Leu	Leu	His	Glu 285	Phe	Gly	Leu	
Leu	Glu 290	Ser	Pro	Lys	Ala	Leu 295	Glu	Glu	Ala	Pro	Trp 300	Pro	Pro	Pro	Glu	
Gly 305	Ala	Phe	Val	Gly	Phe 310	Val	Leu	Ser	Arg	Lys 315	Glu	Pro	Met	Trp	Ala 320	
Asp	Leu	Leu	Ala	Leu 325	Ala	Ala	Ala	Arg	Gly 330	Gly	Arg	Val	His	Arg 335	Ala	
Ala	Asp	Pro	Leu 340	Ala	Gly	Leu	Lys	Asp 345	Leu	Lys	Glu	Val	Arg 350	Gly	Leu	
Leu	Ala	Lys 355	Asp	Leu	Ala	Val	Leu 360	Ala	Ser	Arg	Glu	Gly 365	Leu	Asp	Leu	
Val	Pro 370	Gly	Asp	Asp	Pro	Met 375	Leu	Leu	Ala	Tyr	Leu 380	Leu	Asp	Pro	Ser	
Asn 385	Thr	Thr	Pro	Glu	Gly 390	Val	Ala	Arg	Arg	Tyr 395	Gly	Gly	Glu	Trp	Thr 400	
Glu	Asp	Ala	Ala	His 405	Arg	Ala	Leu	Leu	Ser 410	Glu	Arg	Leu	His	Arg 415	Asn	
Leu	Leu	Lys	Arg 420	Leu	Glu	Gly	Glu	Glu 425	Lys	Leu	Leu	Trp	Leu 430	Tyr	His	
Glu	Val	Glu 435	Lys	Pro	Leu	Ser	Arg 440	Val	Leu	Ala	His	Met 445	Glu	Ala	Thr	
Gly	Val 450	Arg	Arg	Asp	Val	Ala 455	Tyr	Leu	Gln	Ala	Leu 460	Ser	Leu	Glu	Leu	
Ala 465	Glu	Glu	Ile	Arg 470	Arg	Leu	Glu	Glu	Glu	Val 475	Phe	Arg	Leu	Ala	Gly 480	
His	Pro	Phe	Asn 485	Leu	Asn	Ser	Arg	Asp	Gln 490	Leu	Glu	Arg	Val	Leu 495	Phe	
Asp	Glu	Leu	Arg 500	Leu	Pro	Ala	Leu	Gly 505	Lys	Thr	Gln	Lys	Thr 510	Gly	Lys	
Arg	Ser	Thr 515	Ser	Ala	Ala	Val	Leu 520	Glu	Ala	Leu	Arg	Glu 525	Ala	His	Pro	
Ile	Val 530	Glu	Lys	Ile	Leu	Gln 535	His	Arg	Glu	Leu	Thr 540	Lys	Leu	Lys	Asn	

[illegible]

```
<210> 98
<211> 2499
<212> DNA
<213> Artificial Sequence
```

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 98

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
cgggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgccg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg cttcttgaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccttttcc gagaggctct tcgccaacct gtgggggagg 1260
cttgaggggg aggagaggct cttttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaagg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacctctg ccggacctca tccaccccag gacgggcccgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740

aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcgtg 1800
 gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgctctc 1860
 gccacctct cccggggacga aaacctgatc agggctcttc aggaggggaa ggacatccac 1920
 acccagaccg caagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc 2040
 caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
 ttccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
 gtggaaaccc tcttcggaag aaggcgctac gtgcccgaac tcaacgcccg ggtgaagagc 2220
 gtcagggagg ccgaggagcg catggccttc aacatgcccg tccagggcac cgccgcccac 2280
 ctcatgaagc tcgccatggt gaagctcttc ccccgctcc gggagatggg ggcccgcacg 2340
 ctctccagg tccacgacga gtcctcctg gaggcccccc aagcgcgggc cgaggaggtg 2400
 gcggttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggaggtg 2460
 gaggtgggga tgggggagga ctggctttcc gccagggt 2499

<210> 99

<211> 833

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 99

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75				80	
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
			85					90						95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				

Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu	145	150	155	160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	165	170	175	
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	180	185	190	
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu	195	200	205	
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg	210	215	220	
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu	225	230	235	240
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	245	250	255	
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	260	265	270	
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	275	280	285	
Leu	Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	290	295	300	
Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	305	310	315	320
Asp	Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	325	330	335	
Pro	Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu	340	345	350	
Leu	Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	355	360	365	
Pro	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	370	375	380	
Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	385	390	395	400
Glu	Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	405	410	415	
Leu	Trp	Gly	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	420	425	430	
Glu	Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	435	440	445	
Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	450	455	460	
Ala	Glu	Glu	Ile	Ala	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly	465	470	475	480

His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe
				485					490					495	
Asp	Glu	Leu	Gly	Leu	Pro	Ala	Ile	Gly	Lys	Thr	Glu	Lys	Thr	Gly	Lys
			500					505					510		
Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro
		515					520					525			
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser
	530					535					540				
Thr	Tyr	Ile	Asp	Pro	Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg
545					550					555					560
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser
				565					570					575	
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly
			580					585					590		
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Val	Ala	Glu	Ala	Gly	Trp	Ala	Leu	Val
		595					600					605			
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser
	610					615					620				
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Lys	Asp	Ile	His
625					630				635						640
Thr	Gln	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Pro	Glu	Ala	Val	Asp
				645					650					655	
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Val	Asn	Phe	Gly	Val	Leu	Tyr
			660					665					670		
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu
		675					680					685			
Glu	Ala	Val	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val
	690					695					700				
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Lys	Arg	Gly	Tyr
705					710					715					720
Val	Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	Leu	Asn	Ala
				725					730					735	
Arg	Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	Phe	Asn	Met
			740					745				750			
Pro	Val	Gln	Gly	Thr	Ala	Ala	Asp	Leu	Met	Lys	Leu	Ala	Met	Val	Lys
		755					760					765			
Leu	Phe	Pro	Arg	Leu	Arg	Glu	Met	Gly	Ala	Arg	Met	Leu	Leu	Gln	Val
	770					775					780				
His	Asp	Glu	Leu	Leu	Leu	Glu	Ala	Pro	Gln	Ala	Arg	Ala	Glu	Glu	Val
785					790					795					800
Ala	Ala	Leu	Ala	Lys	Glu	Ala	Met	Glu	Lys	Ala	Tyr	Pro	Leu	Ala	Val
				805					810					815	

Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Gly

<210> 100
<211> 2499
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 100
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc ccccacgccg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccgcc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatcccg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgccg gggcttctcg ccaaagacct gagcgttctg 1080
gcctgagggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccgcc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccttttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380

```

tccctggagg tggccgagga gatcgcccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcc tggcaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacggggccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagacc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggtatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacgacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggagg catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag 2499

```

<210> 101

<211> 833

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 101

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln	85	90	95
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu	100	105	110
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys	115	120	125
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys	130	135	140
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu	145	150	155
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	165	170	175
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	180	185	190
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu	195	200	205
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg	210	215	220
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu	225	230	235
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	245	250	255
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	260	265	270
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	275	280	285
Leu	Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	290	295	300
Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	305	310	315
Asp	Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	325	330	335
Pro	Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu	340	345	350
Leu	Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	355	360	365
Pro	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	370	375	380
Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	385	390	395
Glu	Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	405	410	415

1. *Chlorophyll a* (Chl *a*) is the primary photosynthetic pigment in most plants and algae. It is a green pigment that absorbs light energy in the blue and red regions of the visible spectrum. Chl *a* is essential for the light-dependent reactions of photosynthesis, where it converts light energy into chemical energy.

2. *Chlorophyll b* (Chl *b*) is an accessory pigment found in green plants and algae. It is a yellow-green pigment that absorbs light energy in the blue and orange regions of the visible spectrum. Chl *b* transfers the absorbed energy to Chl *a* for use in photosynthesis.

3. *Carotenoids* are a group of pigments that include carotenes and xanthophylls. They are responsible for the yellow, orange, and red colors seen in autumn foliage. Carotenoids absorb light energy in the blue and green regions of the visible spectrum and transfer the energy to Chl *a*. They also play a role in protecting the photosynthetic apparatus from damage by reactive oxygen species.

4. *Xanthophylls* are a subclass of carotenoids that are primarily responsible for the yellow color of autumn leaves. They absorb light energy in the blue and green regions of the visible spectrum and transfer the energy to Chl *a*. Xanthophylls also play a role in the photoprotection of the photosynthetic apparatus.

5. *Anthocyanins* are water-soluble pigments that are responsible for the red, purple, and blue colors seen in many autumn leaves. They are not directly involved in photosynthesis but are produced by the plant in response to environmental factors such as low temperatures and high light intensity.

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu

<210> 102
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 102
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggcct agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020

aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcggttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg ggcaccgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
cttgaggggg aggagagggt cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
tccctggagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
cttcccgctt tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
aagctcaaga acacctacgt ggacccccct ccaagcctcg tccacccgag gacggggccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
aacctgcaga acatccccgt ccgcaccccc ttggggccaga ggatccgccg ggcttctatc 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgct ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcgggggt ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggtatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gaggcccca aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

<210> 103

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 103

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Introduction

Leu	Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	355	360	365
Pro	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	370	375	380
Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	385	390	400
Glu	Glu	Ala	Gly	His	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	405	410	415
Leu	Trp	Gly	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	420	425	430
Glu	Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	435	440	445
Gly	Val	Arg	Arg	Asp	Val	Ala	Tyr	Leu	Gln	Ala	Leu	Ser	Leu	Glu	Leu	450	455	460
Ala	Glu	Glu	Ile	Arg	Arg	Leu	Glu	Glu	Glu	Val	Phe	Arg	Leu	Ala	Gly	465	470	475
His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe	485	490	495
Asp	Glu	Leu	Arg	Leu	Pro	Ala	Leu	Gly	Lys	Thr	Gln	Lys	Thr	Gly	Lys	500	505	510
Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro	515	520	525
Ile	Val	Glu	Lys	Ile	Leu	Gln	His	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Asn	530	535	540
Thr	Tyr	Val	Asp	Pro	Leu	Pro	Ser	Leu	Val	His	Pro	Arg	Thr	Gly	Arg	545	550	555
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	565	570	575
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly	580	585	590
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val	595	600	605
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser	610	615	620
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His	625	630	635
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp	645	650	655
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr	660	665	670
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu	675	680	685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 104
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 104
 gaggaggcgg ggcaccgggc cgccctt

27

<210> 105
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 105
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggcctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300

aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcac 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgcg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgtcctc cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctcc atcggaacct gtgggggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
tccttgagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
cttccgcct tggggaagac gcaaaagaca ggcaagcgt ccaccagcgc cgcggtgctg 1560
gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
aagctcaaga acacctacgt ggacccccct ccaagcctcg tccacccgag gacgggcccgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
aacctgcaga acatccccgt ccgcaccccc ttggggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcgggca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160

gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccc tccagggcac cgccgccgac 2280
 ctcatagaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 106

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 106

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
65					70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
145					150					155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
			165						170					175	
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp
		180						185					190		
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu
	195						200					205			
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg

210	215	220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu 225 230 235 240		
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu 245 250 255		
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala 260 265 270		
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu 275 280 285		
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu 290 295 300		
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala 305 310 315 320		
Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala 325 330 335		
Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu 340 345 350		
Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu 355 360 365		
Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser 370 375 380		
Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr 385 390 395 400		
Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu His Arg Asn 405 410 415		
Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg 420 425 430		
Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr 435 440 445		
Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu 450 455 460		
Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly 465 470 475 480		
His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe 485 490 495		
Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly Lys 500 505 510		
Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro 515 520 525		
Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn 530 535 540		

Thr	Tyr	Val	Asp	Pro	Leu	Pro	Ser	Leu	Val	His	Pro	Arg	Thr	Gly	Arg	
545					550					555					560	
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	
				565					570					575		
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly	
			580					585					590			
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val	
		595					600					605				
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser	
	610					615					620					
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His	
625					630					635				640		
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp	
				645					650					655		
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr	
			660					665					670			
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu	
		675					680					685				
Glu	Ala	Gln	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val	
	690					695					700					
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr	
705					710					715					720	
Val	Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	Leu	Glu	Ala	
				725					730					735		
Arg	Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	Phe	Asn	Met	
			740					745					750			
Pro	Val	Gln	Gly	Thr	Ala	Ala	Asp	Leu	Met	Lys	Leu	Ala	Met	Val	Lys	
		755					760					765				
Leu	Phe	Pro	Arg	Leu	Glu	Glu	Met	Gly	Ala	Arg	Met	Leu	Leu	Gln	Val	
	770					775					780					
His	Asn	Glu	Leu	Val	Leu	Glu	Ala	Pro	Lys	Glu	Arg	Ala	Glu	Ala	Val	
785					790					795					800	
Ala	Arg	Leu	Ala	Lys	Glu	Val	Met	Glu	Gly	Val	Tyr	Pro	Leu	Ala	Val	
			805						810					815		
Pro	Leu	Glu	Val	Glu	Val	Gly	Ile	Gly	Glu	Asp	Trp	Leu	Ser	Ala	Lys	
		820						825					830			
Glu	His	His	His	His	His	His	His									
		835														

<210> 107

<211> 36

<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 107
ctttccgaga ggctccatcg gaacctgtgg gggagg 36

<210> 108
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 108
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgccg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccagccg ggagaggctt agggcctttc tggagaggct tgagtctggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccaggggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccttttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380

tccctggagc ttgcgaggga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
 ccccccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
 cttccccgct tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
 gagggccctac gggaggccca ccccatcgctg gagaagatcc tccagcaccg ggagctcacc 1620
 aagctcaaga acacctacgt ggacccccctc ccaagcctcg tccacccgag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
 aacctgcaga acatccccgt ccgcaccccc ttggggccaga ggatccgccg ggctttcatc 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cgggcgggca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcggac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggtcctc gaggcccca aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

<210> 109

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 109

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80

[illegible]

Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu
 450 455 460
 Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn
 530 535 540
 Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His
835

<210> 110
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 110
ctcttcgccca acctgcttaa gaggcttgag ggggag 36

<210> 111
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 111
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcac 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660

aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccattgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccgcc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtgggggagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttcccg 1320
 gtcttgcccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
 tccctggagc ttgcggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
 cttcccgctt tggggaagac gcaaaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
 gaggccttac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
 aagctcaaga acacctacgt ggacccccct ccaagcctcg tccacccgag gacgggccgc 1680
 ctccacacc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
 aacctgcaga acatccccgt ccgcaccccc ttggggccaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgct ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcgggggtc ctctacggca tgtcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagacct tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggg gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 112
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 112
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu
 450 455 460
 Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn
 530 535 540
 Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 113

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 113

aggccccttt cccgggtcct ggcccat

27

<210> 114

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 114

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtott tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
gcgggggctaa aggacctcaa ggaggtccgg ggccctcctc ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccattgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc cacgggggtg cgccgggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860

gccacacctt cggcgacga gaacctgac cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgct cccggggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca cgcctctcc 2040
 caggagctag ccacccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcttg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcataagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggtcttc gaggcccaa aagagagggc ggaggccgtg 2400
 gccgggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 115

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 115

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75				80	
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
145					150					155					160

1. *Phragmites australis* (Cav.) Trin. ex Steud. (Common reed)
 2. *Scirpus atrovirens* (L.) Link. (Black bog rush)
 3. *Scirpus cespitosus* (L.) Link. (Bog rush)
 4. *Scirpus eriopodus* (L.) Link. (Bog rush)
 5. *Scirpus holosericeus* (L.) Link. (Bog rush)
 6. *Scirpus lacustris* (L.) Link. (Bog rush)
 7. *Scirpus setaceus* (L.) Link. (Bog rush)
 8. *Scirpus tabernaemontani* (L.) Link. (Bog rush)
 9. *Scirpus torreyana* (L.) Link. (Bog rush)
 10. *Scirpus validus* (L.) Link. (Bog rush)
 11. *Scirpus virgatus* (L.) Link. (Bog rush)
 12. *Scirpus yagara* (L.) Link. (Bog rush)
 13. *Scirpus yagara* (L.) Link. (Bog rush)
 14. *Scirpus yagara* (L.) Link. (Bog rush)
 15. *Scirpus yagara* (L.) Link. (Bog rush)
 16. *Scirpus yagara* (L.) Link. (Bog rush)
 17. *Scirpus yagara* (L.) Link. (Bog rush)
 18. *Scirpus yagara* (L.) Link. (Bog rush)
 19. *Scirpus yagara* (L.) Link. (Bog rush)
 20. *Scirpus yagara* (L.) Link. (Bog rush)

Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 116
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 116
acgggggtgc gccgggacgt ggcctat 27

<210> 117
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 117
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaa gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcac 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtggcg accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggcctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccattgctcct cgcctacctc 1140

ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct ccaggccttg 1380
tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccggca tgggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggcca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac 1920
acggagaccg ccagctggat gttcggcgct ccccgaggagg ccgtggaccc cctgatgcgc 1980
cgggcgggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 118

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 118

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	

Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
325 330 335

Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
340 345 350

Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
355 360 365

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 119
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 119
gtggcctatc tccaggcctt gtccctg

27

<210> 120
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 120
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcttacggg 240
gggtacaagg cgggcccggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatacctc 420

accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggagggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg 1020
gcgggggctaa aggacctcaa ggaggtccgg ggccctcctc ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgccctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc cacgggggtg cgccctggacg tggcctatct cagggccttg 1380
tccctggagc ttgccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcga tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
tccccaaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccc ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccc tccagggcac cgccgcccac 2280

ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 121

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 121

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240

[illegible]

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 122

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 122

ttgtccctgg agcttgccga ggagatc

<210> 123
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 123
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggccctacggg 240
 ggggtacaagg cgggccgggc ccccaacgcc gaggactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgccc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggett agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg cttcttgaa agcccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgcc ggggacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
 gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cttttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtcttgggcc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatccgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcga tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560

gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg cgggacctca tccaccccag gacgggccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccc cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct cgggcgacga gaacctgatc cgggtcttcc aggaggggagc ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag cctgggaccc cctgatgcgc 1980
cgggcgggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcttg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgcgcgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cgcccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 124

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 124

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
		20						25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50				55					60					
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70				75					80	
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
			85					90						95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
		100						105					110		

Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	450	455	460	
Ala	Glu	Glu	Ile	Arg	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly	465	470	475	480
His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe	485	490	495	
Asp	Glu	Leu	Gly	Leu	Pro	Ala	Ile	Gly	Lys	Thr	Glu	Lys	Thr	Gly	Lys	500	505	510	
Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro	515	520	525	
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser	530	535	540	
Thr	Tyr	Ile	Asp	Pro	Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg	545	550	555	560
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	565	570	575	
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly	580	585	590	
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val	595	600	605	
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser	610	615	620	
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His	625	630	635	640
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp	645	650	655	
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr	660	665	670	
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu	675	680	685	
Glu	Ala	Gln	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val	690	695	700	
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr	705	710	715	720
Val	Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	Leu	Glu	Ala	725	730	735	
Arg	Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	Phe	Asn	Met	740	745	750	
Pro	Val	Gln	Gly	Thr	Ala	Ala	Asp	Leu	Met	Lys	Leu	Ala	Met	Val	Lys	755	760	765	
Leu	Phe	Pro	Arg	Leu	Glu	Glu	Met	Gly	Ala	Arg	Met	Leu	Leu	Gln	Val	770	775	780	

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 125
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 125
gccgaggaga tccgccgcct cgaggcc 27

<210> 126
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 126
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtcct tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc cctggagggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840

```

agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgcgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
gcggggctaa aggacctcaa ggaggtccgg ggccctctcg ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgcg ctcgaggagg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcc a tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgate cgggtcttcc aggagggggc ggacatccac 1920
acgggagacc ccagctggat gttcggcgtc cccgggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgcgcgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggctctc gaggcccaaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

```

```

<210> 127
<211> 839
<212> PRT
<213> Artificial Sequence

```

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 127

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 128
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 128
 gccgcctcg aggaggaggt cttccgc

27

<210> 129
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 129
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcttacggg 240
 ggggtacaagg cgggccggggc ccccacgccg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tcacgtcct ccaccccgag 480
 ggggtacctca tcaccccgcc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgccc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agaccccttg 1020
 gcgggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgcc gggaagcacc ccatgctcct cgctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcgcc gctacggggg ggagtggacg 1200
 gaggacgccg cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtcttgggcc atatggaggc cacgggggtg cgctggagc tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaagg 1500
 cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccccttg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccctcatc 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920

acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccagggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 130

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 130

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
			85					90						95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
		100						105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
	145				150					155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
			165					170						175	

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
 355 360 365
 Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Lys Leu Leu Trp Leu Tyr His
 420 425 430
 Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys
 500 505 510

<210> 131
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 131
tttgacgagc taaggcttcc cgccatc

27

<210> 132
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 132
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggcttacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtcct tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggg cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccgcc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggcctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgcgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
gcgggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcgcgcc gctacggggg ggagtggacg 1200

Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile	50	55	60
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly	65	70	75
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln	85	90	95
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu	100	105	110
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys	115	120	125
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys	130	135	140
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu	145	150	155
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	165	170	175
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	180	185	190
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu	195	200	205
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg	210	215	220
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu	225	230	235
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	245	250	255
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	260	265	270
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	275	280	285
Leu	Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	290	295	300
Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	305	310	315
Asp	Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	325	330	335
Ala	Asp	Pro	Leu	Ala	Gly	Leu	Lys	Asp	Leu	Lys	Glu	Val	Arg	Gly	Leu	340	345	350
Leu	Ala	Lys	Asp	Leu	Ala	Val	Leu	Ala	Ser	Arg	Glu	Gly	Leu	Asp	Leu	355	360	365
Val	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	370	375	380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His
 420 425 430
 Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Gln Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His
835

<210> 134
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 134
atcgccaaga cgcaaaagac cggcaag

27

<210> 135
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 135
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480

ggggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggagggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgcccg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcggggctaa aggacctcaa ggaggtccgg ggccctctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
 gaggacgcgc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtcttgcccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgcg ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcga tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagcaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgc ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcccga agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 tcccccaagg tgcgggcttg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggg gaagctcttc ccagggtg aggaaatggg ggccaggatg 2340

ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 136
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 136
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 137

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 137

aagatcctgc agcaccggga gctcacc

27

<210> 138
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 138

```

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc ccccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg cttcttgga agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
gcggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccg ccacccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcga tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560

```

gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga acacctacat tgaccccttg ccggacctca tccaccccag gacggggccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccc cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 139

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 139

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
		20						25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105						110	

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 140

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 140

accaagctga agaacaccta cattgac

27

<210> 141

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 141

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag cgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc ccccacgccg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcacctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840

agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgcccg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcggggctaa aggacctcaa ggaggtccgg ggccctctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
 gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtccctggccc atatggaggg cacgggggtg cgccctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac ggagaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
 gaggcctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacgt ggacccttg ccggacctca tccaccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccagggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 142
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 142

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 143
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 143
 aagagcacct acgtggaccc cttgccg

27

<210> 144
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 144
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc ccccacgccg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgccc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtctggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
gcgggggctaa aggacctcaa ggaggtccgg ggcctcctcg ccaaggacct cgccgtcttg 1080
gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
ctcgaggggg agggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
gtcctggccc atatggaggc cacgggggtg cgccctggac tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac ggagaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggcca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacctcttg ccgagcctcg tccaccccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgc ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac 1920

acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
 caggagctag ccattccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcggggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccagggtgg aggaaatggg ggccaggatg 2340
 ctccttcagg tccacaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 145

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 145

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175

<210> 146
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 146
attgaccctt tgccgagcct cgtccacccc aggacgggc 39

<210> 147
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 147
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag cgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgccc gaggactttc cccggcaact cggcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccgcc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggc caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgccc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggcct agggcctttc tggagaggct tgagtcttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggc aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcggcg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200

gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
tccttgagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggcgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggaccc cctgatgcgc 1980
cgggcgccca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagacc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcttc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaggagc accaccacca ccaccac 2517

<210> 148

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 148

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	

Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
		20						25					30		

Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			

[illegible]

Figure 1. The effect of the concentration of the *Agrobacterium* suspension on the transformation efficiency of *Agrobacterium* strains. The *Agrobacterium* strains were grown in YEA medium at 28°C for 24 h. The cell concentration was adjusted to 10⁸ cells/ml. The cells were then mixed with the plant tissue and the transformation efficiency was determined. The results are shown as the mean ± SD of three independent experiments. The asterisk indicates a significant difference (*p* < 0.05) between the two strains.

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 149
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 149
atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgccc gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatoctc 420
accgcccaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccgggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840


```

agcctcctcc acgagttcgg ccttcttgaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgcgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttgggagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
tccttgaggg tggccgagga gatcgccgcg ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac gcaaaagacc ggcaagcgtc ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacctcttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgc ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
cgggcgggcca agaccatcaa cttcggggtc ctctacggca tgtcgggcca ccgcctctcc 2040
caggagctag ccacccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttcccaagg tgccggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggg gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tccacaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtgg ccaaggagg catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

```

```

<210> 150
<211> 839
<212> PRT
<213> Artificial Sequence

```

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 150

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala

1. *Chlorophyll a* (Chl *a*) is the primary photosynthetic pigment in most plants and algae. It is a green pigment that absorbs light energy in the blue and red regions of the visible spectrum. Chl *a* is essential for the light-dependent reactions of photosynthesis, where it converts light energy into chemical energy.

2. *Chlorophyll b* (Chl *b*) is an accessory pigment found in green plants and algae. It absorbs light energy in the blue and orange-red regions of the visible spectrum. Chl *b* transfers the absorbed energy to Chl *a* for use in photosynthesis.

3. *Carotenoids* are a group of pigments that include carotenes and xanthophylls. They absorb light energy in the blue and green regions of the visible spectrum. Carotenoids transfer energy to Chl *a* and also play a role in protecting the photosynthetic apparatus from damage by excess light energy.

4. *Xanthophylls* are a subset of carotenoids that include pigments like lutein and zeaxanthin. They absorb light energy in the blue and green regions of the visible spectrum. Xanthophylls are involved in the xanthophyll cycle, which helps regulate the light-harvesting capacity of the photosynthetic apparatus under varying light conditions.

5. *Anthocyanins* are water-soluble pigments that give plants red, purple, and blue colors. They are not directly involved in photosynthesis but can play a role in protecting plants from environmental stressors like UV radiation and herbivory.

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 His Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 151
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 151
 ttcgccaacc tgcttgggag gcttgagggg gag

33

<210> 152
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 152
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtcct tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccgggc ccccacgccg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatc ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accggggccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcg gggttctctg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgcc gccgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gtggaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc ctttccgct 1320
 gtctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaagg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccttgg ccggacctca tccacccag gacgggcccgc 1680
 ctccacacc gcttaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920

1. *Chlorophyll a* (Chl *a*) is the primary photosynthetic pigment in most plants and algae. It is a green pigment that absorbs light energy in the blue and red regions of the visible spectrum. Chl *a* is essential for the light-dependent reactions of photosynthesis, where it converts light energy into chemical energy.

2. *Chlorophyll b* (Chl *b*) is an accessory pigment found in green plants and algae. It absorbs light energy in the blue and orange-red regions of the visible spectrum. Chl *b* transfers the absorbed energy to Chl *a* for use in photosynthesis.

3. *Carotenoids* are a group of pigments that include carotenes and xanthophylls. They absorb light energy in the blue and green regions of the visible spectrum. Carotenoids transfer energy to Chl *a* and also play a role in protecting the photosynthetic apparatus from damage by excess light energy.

4. *Xanthophylls* are a subset of carotenoids that include pigments like lutein and zeaxanthin. They absorb light energy in the blue and green regions of the visible spectrum. Xanthophylls are involved in the xanthophyll cycle, which helps regulate the light absorption capacity of the photosynthetic apparatus under varying light conditions.

5. *Anthocyanins* are water-soluble pigments that give plants red, purple, and blue colors. They are not directly involved in photosynthesis but can play a role in protecting plants from environmental stressors like UV radiation and herbivory.

<211> 839

<213> Artificial Sequence

<223> Description of Artificial Sequence: Synthetic

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

[illegible]

Arg	Ser	Thr 515	Ser	Ala	Ala	Val	Leu 520	Glu	Ala	Leu	Arg	Glu 525	Ala	His	Pro
Ile	Val 530	Glu	Lys	Ile	Leu	Gln 535	Tyr	Arg	Glu	Leu	Thr 540	Lys	Leu	Lys	Ser
Thr 545	Tyr	Ile	Asp	Pro	Leu 550	Pro	Asp	Leu	Ile	His 555	Pro	Arg	Thr	Gly	Arg 560
Leu	His	Thr	Arg	Phe 565	Asn	Gln	Thr	Ala	Thr 570	Ala	Thr	Gly	Arg	Leu 575	Ser
Ser	Ser	Asp	Pro 580	Asn	Leu	Gln	Asn	Ile 585	Pro	Val	Arg	Thr	Pro 590	Leu	Gly
Gln	Arg	Ile 595	Arg	Arg	Ala	Phe	Ile 600	Ala	Glu	Glu	Gly	Trp 605	Leu	Leu	Val
Ala	Leu 610	Asp	Tyr	Ser	Gln	Ile 615	Glu	Leu	Arg	Val	Leu 620	Ala	His	Leu	Ser
Gly 625	Asp	Glu	Asn	Leu	Ile 630	Arg	Val	Phe	Gln	Glu 635	Gly	Arg	Asp	Ile	His 640
Thr	Glu	Thr	Ala	Ser 645	Trp	Met	Phe	Gly	Val 650	Pro	Arg	Glu	Ala	Val 655	Asp
Pro	Leu	Met	Arg 660	Arg	Ala	Ala	Lys	Thr 665	Ile	Asn	Phe	Gly	Val 670	Leu	Tyr
Gly	Met	Ser 675	Ala	His	Arg	Leu	Ser 680	Gln	Glu	Leu	Ala	Ile 685	Pro	Tyr	Glu
Glu	Ala 690	Gln	Ala	Phe	Ile	Glu 695	Arg	Tyr	Phe	Gln	Ser 700	Phe	Pro	Lys	Val
Arg 705	Ala	Trp	Ile	Glu	Lys 710	Thr	Leu	Glu	Glu	Gly 715	Arg	Arg	Arg	Gly	Tyr 720
Val	Glu	Thr	Leu	Phe 725	Gly	Arg	Arg	Arg	Tyr 730	Val	Pro	Asp	Leu	Glu 735	Ala
Arg	Val	Lys	Ser 740	Val	Arg	Glu	Ala	Ala 745	Glu	Arg	Met	Ala	Phe 750	Asn	Met
Pro	Val	Gln 755	Gly	Thr	Ala	Ala	Asp 760	Leu	Met	Lys	Leu	Ala 765	Met	Val	Lys
Leu	Phe 770	Pro	Arg	Leu	Glu	Glu 775	Met	Gly	Ala	Arg	Met 780	Leu	Leu	Gln	Val
His 785	Asn	Glu	Leu	Val	Leu 790	Glu	Ala	Pro	Lys	Glu 795	Arg	Ala	Glu	Ala	Val 800
Ala	Arg	Leu	Ala	Lys 805	Glu	Val	Met	Glu	Gly 810	Val	Tyr	Pro	Leu	Ala 815	Val
Pro	Leu	Glu	Val 820	Glu	Val	Gly	Ile	Gly 825	Glu	Asp	Trp	Leu	Ser 830	Ala	Lys
Glu	His 835	His	His	His	His	His									

<210> 154
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 154
 ctcttcgccca acctgtggaa gaggttgag ggg

33

<210> 155
 <211> 835
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 155
 Met Glu Phe Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
 1 5 10 15
 Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
 20 25 30
 Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
 35 40 45
 Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
 65 70 75 80
 Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg Leu
 100 105 110
 Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
 115 120 125
 Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
 130 135 140
 Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
 145 150 155 160
 Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
 165 170 175
 Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
 180 185 190
 Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
 195 200 205
 Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg

Ser Thr Tyr Leu Asp Pro Leu Pro Arg Leu Val His Pro Arg Thr Gly
 545 550 555 560
 Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
 565 570 575
 Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
 580 585 590
 Gly Gln Arg Ile Arg Lys Ala Phe Val Ala Glu Glu Gly Trp Leu Leu
 595 600 605
 Leu Ala Ala Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
 610 615 620
 Ser Gly Asp Glu Asn Leu Lys Arg Val Phe Arg Glu Gly Lys Asp Ile
 625 630 635 640
 His Thr Glu Thr Ala Ala Trp Met Phe Gly Leu Asp Pro Ala Leu Val
 645 650 655
 Asp Pro Lys Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
 660 665 670
 Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Gly Ile Asp Tyr
 675 680 685
 Lys Glu Ala Glu Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
 690 695 700
 Val Arg Ala Trp Ile Glu Arg Thr Leu Glu Glu Gly Arg Thr Arg Gly
 705 710 715 720
 Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Ala
 725 730 735
 Ser Arg Val Arg Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
 740 745 750
 Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Ile Ala Met Val
 755 760 765
 Lys Leu Phe Pro Arg Leu Lys Pro Leu Gly Ala His Leu Leu Leu Gln
 770 775 780
 Val His Asp Glu Leu Val Leu Glu Val Pro Glu Asp Arg Ala Glu Glu
 785 790 795 800
 Ala Lys Ala Leu Val Lys Glu Val Met Glu Asn Ala Tyr Pro Leu Asp
 805 810 815
 Val Pro Leu Glu Val Glu Val Gly Val Gly Arg Asp Trp Leu Glu Ala
 820 825 830
 Lys Gln Asp
 835

<210> 156
 <211> 2526
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 156

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcgg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccaggg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggctcc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggcgcc 420
atcctcaccg ccgaccgcga cctctacca ctcgtctccg accgcgtcgc cgtcctccac 480
cccaggggcc acctcatcac ccgagtggtg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttcgcgc cctcgtgggg gacccctccg acaacctccc cggggtaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcgggg gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctctggagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tctctcccc ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcagggacg gccgggtgca ccgggcagca 1020
gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctcgtgcccg gggacgacct catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgcg ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgagg ggaaaagccc 1320
ctctccccgg tctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgcccgc tcgaggagga ggtcttccgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
acgggccgcc tcacacccc cttcaaccag acggccacgg ccacggggag gcttagtagc 1740

```

tccgacccca acctgcagaa catccccgtc cgcaccccct tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
cgcgctctcg cccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatcgcca cccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
ctgatgcgcc gggcgggcaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg 2220
gtgaagagcg tcaggagggc cgcgagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcatgc tcctccaggt ccacaacgag ctctcctgag agggccccca agcgcgggcc 2400
gaggaggtgg cggttttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
caccac

```

```

<210> 157
<211> 842
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic

```

```

<400> 157
Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
  1             5             10             15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
          20             25             30
Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
          35             40             45
Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
          50             55             60
Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
          65             70             75             80
Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
          85             90             95
Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
          100            105            110
Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
          115            120            125

```

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
 130 135 140
 Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
 145 150 155 160
 Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
 165 170 175
 Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
 180 185 190
 Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
 195 200 205
 Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
 210 215 220
 Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
 225 230 235 240
 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
 245 250 255
 Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val
 325 330 335
 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
 355 360 365
 Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
 420 425 430
 Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
 435 440 445
 Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
 450 455 460

Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
 530 535 540
 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575
 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
 595 600 605
 Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
 625 630 635 640
 Asp Ile Ala Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val His Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
 785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 158

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 158

gcttgccggtc tgggtggcga tgccttccc etc

33

<210> 159

<211> 2526

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 159

atgaattccg aggcgatgct tccgctcttt gaaccctaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggtctcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gcccaaggccc cctccttccg ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccagag acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggctacgag 360
gcgacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgcga cctctaccaa ctctgtctcc accgcgtcgc cgtcctccac 480
cccagggggc acctcatcac ccgggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc agggggcgga gcccgaccgg gaggggctta gggccttccg ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960


```

atgtgggagg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca 1020
gaccccttgg cgggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgtcccg gggacgaccc catgctcctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
gagtggacgg aggacgccgc ccaccgggcc ctctctcgcg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accgggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttcccgctt gggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg 1680
acgggcccgc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tcgacccca acctgcagaa catcccgcgc cgcacccctc tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttggcggttg gtggccctgg actatagcca gatagagctc 1860
cgctcctcg ccacctctc cggggacgaa aacctgatca gggctctcca ggagggaag 1920
gacatccaca ccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
ctgatgcgcc gggcgccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccat 2040
aggctctccc aggagcttgc catccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaagt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaacctt cttcggaaga aggcgctacg tgcccaccc caacgcccgg 2220
gtgaagagcg tcaggagggc cgcggaggcc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcacgc tcctccaggt ccacaacgag ctctcctcgg agggccccc agcgcgggcc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccctc cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
caccac 2526

```

<210> 160

<211> 842

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 160

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160

Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175

Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190

Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205

Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220

Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240

Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
245 250 255

Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
260 265 270

Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
275 280 285

Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
290 295 300

Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
305 310 315 320

Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val

325										330										335									
His	Arg	Ala	Ala	Asp	Pro	Leu	Ala	Gly	Leu	Lys	Asp	Leu	Lys	Glu	Val														
			340					345						350															
Arg	Gly	Leu	Leu	Ala	Lys	Asp	Leu	Ala	Val	Leu	Ala	Ser	Arg	Glu	Gly														
		355					360					365																	
Leu	Asp	Leu	Val	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu														
	370					375					380																		
Asp	Pro	Ser	Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly														
385					390					395					400														
Glu	Trp	Thr	Glu	Asp	Ala	Ala	His	Arg	Ala	Leu	Leu	Ser	Glu	Arg	Leu														
				405					410					415															
His	Arg	Asn	Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Lys	Leu	Leu	Trp														
			420					425					430																
Leu	Tyr	His	Glu	Val	Glu	Lys	Pro	Leu	Ser	Arg	Val	Leu	Ala	His	Met														
	435						440					445																	
Glu	Ala	Thr	Gly	Val	Arg	Arg	Asp	Val	Ala	Tyr	Leu	Gln	Ala	Leu	Ser														
	450					455					460																		
Leu	Glu	Leu	Ala	Glu	Glu	Ile	Arg	Arg	Leu	Glu	Glu	Glu	Val	Phe	Arg														
465					470					475				480															
Leu	Ala	Gly	His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg														
				485					490				495																
Val	Leu	Phe	Asp	Glu	Leu	Arg	Leu	Pro	Ala	Leu	Gly	Lys	Thr	Gln	Lys														
			500					505					510																
Thr	Gly	Lys	Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu														
		515					520					525																	
Ala	His	Pro	Ile	Val	Glu	Lys	Ile	Leu	Gln	His	Arg	Glu	Leu	Thr	Lys														
	530					535				540																			
Leu	Lys	Asn	Thr	Tyr	Val	Asp	Pro	Leu	Pro	Ser	Leu	Val	His	Pro	Arg														
545					550					555				560															
Thr	Gly	Arg	Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly														
				565					570					575															
Arg	Leu	Ser	Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr														
			580					585					590																
Pro	Leu	Gly	Gln	Arg	Ile	Arg	Arg	Ala	Phe	Val	Ala	Glu	Ala	Gly	Trp														
		595					600					605																	
Ala	Leu	Val	Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala														
		610				615					620																		
His	Leu	Ser	Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Lys														
625					630				635					640															
Asp	Ile	His	Thr	Gln	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Pro	Glu														
				645					650					655															

gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcgg ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccc acccccagg acttcccccg gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggtccc cggctacgag 360
 gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
 atcctcaccg ccgaccgcga cctctaccaa ctctgtctcc accgcgtcgc cgtcctccac 480
 cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttcgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
 ggcacgcggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
 gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
 gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
 gagttcggca gcctcctcca cgagttcggc ctctggagg cccccgccc cctggaggag 900
 gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
 atgtgggcgg agcttaaagc cctggccgcc tgcagggcg gccgcgtgca ccgggcagca 1020
 gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgacct catgctctc 1140
 gcctacctcc tggaccttc gaacaccacc cccgagggg tggcgcggcg ctacgggggg 1200
 gagtggacgg aggacgcgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
 cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgagg ggaaaagccc 1320
 ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
 caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
 ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
 gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
 gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
 gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
 acgggcccgc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
 tccgaccca acctgcagaa catccccgtc cgcacccct tgggcccagag gatccgccg 1800
 gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
 cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggcttcca ggaggggaag 1920
 gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccgaggc cgtggacccc 1980

```

ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccac 2040
aggtctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaagg gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgcact caacgcccgg 2220
gtgaagagcg tcagggaggg cgcgagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcatgc tctccaggt cgccaacgag ctctctctgg agggccccca agcgcgggcc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
caccac 2526

```

<210> 163

<211> 842

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 163

```

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
 1             5             10             15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
      20             25             30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
      35             40             45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
      50             55             60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
      65             70             75             80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
      85             90             95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
      100            105            110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
      115            120            125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
      130            135            140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
      145            150            155            160

Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
      165            170            175

```


<210> 164
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 164
caggaggagc tcgttggcga cctggaggag 30

<210> 165
<211> 2526
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 165
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaggccc cctccttcog ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccc acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgccg tgcagggtccc cggctacgag 360
gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgagggtcgc 420
atcctcaccg ccgaccgga cctctacca ctctctccg accgcgtcgc cgtcctccac 480
cccaggggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cgggggtcaag 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccc ccccgagccc 960
atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gacctcttgg cgggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgtcccc gggacgacct catgctcctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200

gagtgagcgg aggcgcgcg ccacggggcc ctctctcgg agaggctcca tcggaacctc 1260

cttaagcgcc tcgaggggga ggagaagctc ctttggtct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accggggtag ggcgggacgt ggcctacctt 1380
caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttcgc 1440
ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac 1500
gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggccttacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctcacca agtcaagaa cacctacgtg gacccccctc caagcctcgt ccacccgagg 1680
acgggcccgc tccacaccgc cttaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgacccca acctgcagaa catccccgtc cgcacccct tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
cgcgtcctcg ccacctctc cggggacgaa aacctgatca gggctctcca ggaggggaag 1920
gacatccaca ccagaccgc aagctggatg ttcggcgtcc ccccgaggc cgtggacccc 1980
ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccac 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tcccgaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg 2220
gtgaagagcg tcaggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgcccacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcatgc tcctccaggt cgccaacgag ctctcctcgg agggccccca agcgcgggcc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaaggggtc ccaccaccac 2520
caccac 2526

<210> 166
<211> 842
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 166
Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
65 70 75 80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
85 90 95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100 105 110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115 120 125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130 135 140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145 150 155 160

Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165 170 175

Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180 185 190

Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
195 200 205

Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210 215 220

Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225 230 235 240

Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
245 250 255

Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
260 265 270

Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
275 280 285

Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
290 295 300

Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
305 310 315 320

Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
325 330 335

His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
340 345 350

Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
355 360 365

Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
 420 425 430
 Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
 435 440 445
 Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
 450 455 460
 Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
 530 535 540
 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575
 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
 595 600 605
 Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
 625 630 635 640
 Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 167
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 167
ggagcgcttg cctgtcttct tcgtcttctt caaggcggga ggcct 45

<210> 168
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 168
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggcttacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420

accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 gggtaacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgccg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cttttggctt taccgggagg tggagaggcc cttttccgct 1320
 gtcttgcccc atatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac gcaaaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
 gaggcctcc gcgaggcca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggcccgc 1680
 ctccacacc gttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgc ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgccggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagacc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280

ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 169

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 169

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
65					70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
	115						120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
145					150					155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
			165					170						175	
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp
		180						185					190		
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu
	195						200					205			
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg
	210					215					220				
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu
225					230					235					240

Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	245	250	255
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	260	265	270
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	275	280	285
Leu	Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	290	295	300
Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	305	310	315
Asp	Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	325	330	335
Pro	Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu	340	345	350
Leu	Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	355	360	365
Pro	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	370	375	380
Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	385	390	395
Glu	Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	405	410	415
Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	420	425	430
Glu	Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	435	440	445
Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	450	455	460
Ala	Glu	Glu	Ile	Ala	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly	465	470	475
His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe	485	490	495
Asp	Glu	Leu	Gly	Leu	Pro	Ala	Ile	Gly	Lys	Thr	Gln	Lys	Thr	Gly	Lys	500	505	510
Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro	515	520	525
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser	530	535	540
Thr	Tyr	Ile	Asp	Pro	Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg	545	550	555
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	565	570	575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 170

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 170

gaggaccagc tcgttgccga cctgaaggag cat

33

<210> 171
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 171
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
 ggggtacctca tcaccccgcc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcctc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccgcc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560

gagggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacggggccgc 1680
ctccacacccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcacccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatcgcc 1920
acggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggaccc cctgatgcgc 1980
cgggcgggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaggagc accaccacca ccaccac 2517

<210> 172

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 172

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
65					70					75				80	
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105						110	

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
450 455 460

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
485 490 495

Asp Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Gln Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile Ala
625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 173
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 173
gaggggcggg acatcgccac ggagaccgcc agc 33

<210> 174
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 174
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac ctccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacg ctccgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgccg gaggaacttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccagctcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagagggt tgagtttgcc 840

```

agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgagg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttcttg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgagg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccttgaggg tggccgagga gatcgccgcg ctcgaggcgg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcggcaagac gcaaaagacc ggcaagcgtt ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacggggcgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt cgccaccccg cttgggcaga ggatccgcgg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc agggggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
cgggcgggca agaccatcaa cttcggggtc ctctacggca tgtcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgccggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggg gaagctcttc ccagggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggtcctc gaggcccca aagagagggc ggaggccgtg 2400
gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

```

```

<210> 175
<211> 839
<212> PRT
<213> Artificial Sequence

```

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 175

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 176

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 176

cagaacatcc ccgtcgccac cccgcttggg cag

33

<210> 177

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 177

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgccc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggcct agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ctttctggaa agcccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggtctctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctccgcgcc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccaggggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgcccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
cttcccgcca tcaagaagac gcaaaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccttctg ccggacctca tccacccag gacgggcccgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920

acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatacctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcggggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccagggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gagtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 178

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 178

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
			35				40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
145				150						155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
			165						170					175	

Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro
		515					520					525			
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser
	530					535					540				
Thr	Tyr	Ile	Asp	Pro	Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg
545					550					555					560
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser
				565					570					575	
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly
			580					585					590		
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val
		595					600					605			
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser
	610					615					620				
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His
625					630					635					640
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp
				645					650					655	
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr
			660					665					670		
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu
		675					680					685			
Glu	Ala	Gln	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val
	690					695					700				
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr
705					710					715					720
Val	Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	Leu	Glu	Ala
				725					730					735	
Arg	Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	Phe	Asn	Met
			740					745					750		
Pro	Val	Gln	Gly	Thr	Ala	Ala	Asp	Leu	Met	Lys	Leu	Ala	Met	Val	Lys
		755					760					765			
Leu	Phe	Pro	Arg	Leu	Glu	Glu	Met	Gly	Ala	Arg	Met	Leu	Leu	Gln	Val
	770					775					780				
Ala	Asn	Glu	Leu	Val	Leu	Glu	Ala	Pro	Lys	Glu	Arg	Ala	Glu	Ala	Val
785					790					795					800
Ala	Arg	Leu	Ala	Lys	Glu	Val	Met	Glu	Gly	Val	Tyr	Pro	Leu	Ala	Val
				805					810					815	
Pro	Leu	Glu	Val	Glu	Val	Gly	Ile	Gly	Glu	Asp	Trp	Leu	Ser	Ala	Lys
			820					825					830		
Glu	His	His	His	His	His	His									
		835													

<210> 179
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 179
gggcttcccg ccatcaagaa gacggagaag acc

33

<210> 180
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 180
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
gggtacaagg cgggccgggc cccacgcccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatectc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggcgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200

gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtccctggccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc cggggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac gcaaaagacc ggcaagcgct ccaccagcg cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacggggccg 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tcggggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggtatggg gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagaggcg ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaggagc accaccacca ccaccac 2517

<210> 181
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 181
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
325 330 335

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
340 345 350

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
355 360 365

Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
370 375 380

Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	
385					390					395					400	
Glu	Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	
				405					410					415		
Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	
			420					425					430			
Glu	Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	
		435					440					445				
Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	
	450					455					460					
Ala	Glu	Glu	Ile	Ala	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly	
465					470					475					480	
His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe	
			485						490					495		
Asp	Glu	Leu	Gly	Leu	Pro	Ala	Ile	Gly	Lys	Thr	Gln	Lys	Thr	Gly	Lys	
			500					505						510		
Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro	
		515					520					525				
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser	
	530					535						540				
Thr	Tyr	Ile	Asp	Pro	Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg	
545					550					555					560	
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	
				565					570					575		
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly	
			580					585					590			
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val	
		595					600					605				
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser	
	610					615					620					
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His	
625					630					635				640		
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp	
				645					650					655		
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr	
			660					665					670			
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu	
		675					680					685				
Glu	Ala	Gln	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val	
	690					695					700					
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr	
705					710					715					720	

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 182
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 182
ctagggcttc ccgccatcaa gaagacgcaa aagaccggc

39

<210> 183
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 183
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggctctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cgagacggag gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480

gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaacc ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgaaggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cttttggctt taccgggagg tggagaggcc cttttccgct 1320
 gtcttgccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccttgagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcca tcggcaagac gcaaaagacc ggcaagcgt ccaccagcgc cgccgtcctg 1560
 gaggcctcc gcgaggccca ccccatcggt gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccccttg ccggacctca tccaccccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctcaga acatccccgt ccgcaccccg cttgggcaga ggatccgcgg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc cccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttcccaagg tgcgggctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340

ctccttcagg tcgccaacga gctggtcctc gaggcccca aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 184
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 184
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Glu Thr Glu Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 185

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 185

ccgggggaaag tcctcctccg tctcggcccg gcccgctt

<210> 186
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 186
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
 ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccgggc cccacgcgcg gaggactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcg accgacctgc ccctggagggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttegg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgcc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct ctttggctt taccgggagg tggagaggcc cttttccgct 1320
 gtcttgcccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg 1380
 tccttgagg tggccgagga gatcgccgc ctgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaggg 1500
 cttcccgcga tcggcaagac gcaaaagacc ggcaagcgct ccaccagcgc cgccgtcctg 1560

gaggccctcc gcgaggccca ccccatcggtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg cgggacctca tccaccccag gacggggccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt cgcacccccg cttgggcaga ggatccgccc ggcttcatc 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcgggcca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggtctg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggag accaccacca ccaccac 2517

<210> 187

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 187

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
		20						25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
		50				55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75				80	
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr	Arg	Leu
			100					105						110	

Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys	
		115					120					125				
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys	
		130				135					140					
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu	
		145			150					155					160	
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	
				165					170					175		
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	
			180					185					190			
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu	
		195					200					205				
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg	
		210				215					220					
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu	
		225			230					235					240	
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	
				245					250					255		
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	
			260					265					270			
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	
		275					280					285				
Leu	Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	
		290				295					300					
Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	
		305			310					315					320	
Asp	Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	
				325					330					335		
Pro	Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu	
			340					345					350			
Leu	Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	
		355					360					365				
Pro	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	
		370				375					380					
Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	
		385			390					395					400	
Glu	Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	
				405					410					415		
Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg	
			420					425					430			
Glu	Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr	
		435					440					445				

Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val	450	455	460
Ala	Glu	Glu	Ile	Ala	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly	465	470	475
His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe	485	490	495
Asp	Glu	Leu	Gly	Leu	Pro	Ala	Ile	Gly	Lys	Thr	Gln	Lys	Thr	Gly	Lys	500	505	510
Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro	515	520	525
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser	530	535	540
Thr	Tyr	Ile	Asp	Pro	Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg	545	550	555
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	565	570	575
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly	580	585	590
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val	595	600	605
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser	610	615	620
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His	625	630	635
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp	645	650	655
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr	660	665	670
Gly	Met	Ser	Ala	His	Arg	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu	675	680	685
Glu	Ala	Gln	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val	690	695	700
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr	705	710	715
Val	Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	Leu	Glu	Ala	725	730	735
Arg	Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	Phe	Asn	Met	740	745	750
Pro	Val	Gln	Gly	Thr	Ala	Ala	Asp	Leu	Met	Lys	Leu	Ala	Met	Val	Lys	755	760	765
Leu	Phe	Pro	Arg	Leu	Glu	Glu	Met	Gly	Ala	Arg	Met	Leu	Leu	Gln	Val	770	775	780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 188
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 188
cgggacctcg aggcgcgtga accccaggag gtccac 36

<210> 189
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 189
atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840

```

agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgagg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
tccttgaggg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
cttcccaagt tgaagaagac gaagaagacc ggtaagcgt ccaccagcg cgcctcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacggggcgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgctcgccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagacct tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggg gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
gccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

```

```

<210> 190
<211> 839
<212> PRT
<213> Artificial Sequence

```

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 190

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 191
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 191
 ctccctccacg agttcggc

18

<210> 192
 <211> 48
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 192
 accggtcttc ttcgtcttct tcaacttggg aagcctgagc tcgtcaaa

48

<210> 193
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 193
aagacgaaga agaccggtaa gcgctccacc agc 33

<210> 194
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 194
gtcgactcta gatcagtggg ggtgggtggg gtgcttgccc gcccggcgca tc 52

<210> 195
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<221> modified_base
<222> (19)..(42)
<223> The bases in these positions within this primer
are 91% of the base shown and 3% each of the other
3 nucleotides.

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 195
ggagcgctta ccggtctttt gcgtcttctt gatcttgcca agccttagct cgtcaaagag 60

<210> 196
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 196
ctcctcacg agttcggc 18

<210> 197
<211> 60
<212> DNA
<213> Artificial Sequence

<220>

<221> modified base
 <222> (19)..(42)
 <223> The bases at these positions within this primer
 are 91% of the base shown and 3% each of the other
 3 nucleotides.

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 197
 caaaagaccg gtaagcgctc caccagcgcc gccgtcctgg aggccctccg cgaggccac 60

<210> 198
 <211> 52
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 198
 gtcgactcta gatcagtggg ggtgggtggg gtgcttggcc gcccgcgca tc 52

<210> 199
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 199
 atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg octaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggcccgggc ccccacgccg gaggactttc cccggcaact cgccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
 aagctctcct gggacctggc caaggtgctc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840

```

agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcgagg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cttcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct ctttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccttgaggg tggccgagga gatcgccgc ctcgaggccg aggtcttcg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaagg tcctctttga cgagctaagg 1500
cttccaaga tcaacaagac gaagaagacc ggtaagcgt ccaccagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgt gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccaccccag gacgggcccgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgcg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
cgggcgggca agaccatcaa ctteggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttcccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
gtggagaccc tcttcggccg ccgccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg 2340
ctccttcagg tcgccaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

```

```

<210> 200
<211> 839
<212> PRT
<213> Artificial Sequence

```

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 200

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240
Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255
Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270
Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285
Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300
Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 201

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 201

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cggctacagg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtcct tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggccgggc cccacgccc gaggaacttc cccggcaact cgccctcctc 300
 aaggagctgg tggacctcct ggggctggcg cgctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420

accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 ggggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
 ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggagggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggccttcc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
 gatcttcttg ccctggccgc cgccaggggc ggccgcgtcc accggggccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccgcc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct 1320
 gtcctggccc atatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgcccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctaagg 1500
 attcccaaga tcaagaagac gcataagacc ggtaagcgct ccaccagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgct gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacctcttg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgcccgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280

ctcatgaagc tggctatggt gaagctcttc ccaggtctgg aggaaatggg ggccaggatg 2340
ctccttcagg tgcgaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 202
<211> 839
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 202
Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
1 5 10 15
Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
20 25 30
Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
35 40 45
Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80
Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140
Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160
Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205
Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220
Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 203

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 203

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcggtggtctt tgacgccaag gccccctcct tccgccacga ggccctacggg 240
 gggtagaagg cggggccgggc cccacgcgg gaggactttc cccggcaact cggccctcatc 300
 aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccag 480
 gggtagctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc tcccgggggt caagggcac 600
 ggggagaaga cggcgaggaa gcttctggag gagggggga gcctggaagc cctcctcaag 660
 aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggaogatctg 720
 aagctctcct gggacctggc caaggcgcg accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ctttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct ctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtcttgggcc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccttgaggg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
 cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacctttg ccggacctca tccaccccag gacgggcccgc 1680
 ctccacaccc gtttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920

acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccattccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcggggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccc ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcccttcagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaggagc accaccacca ccaccac 2517

<210> 204

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 204

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	

Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
		20						25					30		

Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			

Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				

Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
65					70					75					80

Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	

Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Leu	Ala	Arg	Leu
		100						105					110		

Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			

Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				

Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
145					150					155					160

Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
				165					170					175	

Arg Ser Ser Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 205
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 205

```

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgcgcg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgcbc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cttttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgctggagc tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccaccagcgc cgccctcctg 1560

```

gagggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacccttg cgggacctca tccaccccag gacggggccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcacccccg cttgggcaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggag ggacatccac 1920
 acggagaccg ccagctggat gttcggcgct ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcgggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cgcccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggtatggg gaagctcttc ccagggtgg aggaaatggg ggccaggatg 2340
 ctccttcagg tcgccaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 206

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 206

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
 100 105 110

Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
450 455 460

Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
465 470 475 480

His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
485 490 495

Asp Glu Leu Arg Leu Pro Lys Leu Lys Lys Thr Lys Lys Thr Gly Lys
500 505 510

Arg Ser Thr Ser Ala Ala Leu Leu Glu Ala Leu Arg Glu Ala His Pro
515 520 525

Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
530 535 540

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
545 550 555 560

Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
565 570 575

Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
580 585 590

Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
595 600 605

Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
625 630 635 640

Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
645 650 655

Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
660 665 670

Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
690 695 700

Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830

Glu His His His His His His
 835

<210> 207
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 207
 gtcggagggg tccccacga g 21

<210> 208
 <211> 17
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 208
 tgtggaattg tgagcgg 17

<210> 209
 <211> 75
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> modified_base
 <222> (28)..(59)
 <223> The bases in these positions within this primer
 are 91% of the base shown and 3% each of the other
 3 nucleotides.

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 209
 ctcgtggggg acccctccga caacctcccc ggggtcaagg gcatcgggga gaagaccgcc 60
 ctcaagcttc tcaag 75

<210> 210
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 210
gtggcctcca tatgggccag gac

23

<210> 211
<211> 2526
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 211
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaagccc cctccttcgc ccacgaggcc 240
tacgaggcct acaaggcggg gagggccccg acccccaggg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg tttaccgcgc tcgaggctcc cggctacgag 360
gcgagcgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgcga cctctacca aatcgctcgc accgcgtcgc cgtcctccac 480
cccagaggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccg aggggtcagg 600
ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggccacactg 720
gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
gacctcggcc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
gagttcggca gcctcctcca cgagttcggc ctctggagg cccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
atgtggggcg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gacccttggt cgggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gcgctcttgg cctcgaggga ggggctagac ctctgcccc gggacgaccc catgctcctc 1140
gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcggcg ctacgggggg 1200
gagtggacgg aggacggcg ccaccgggct ctctctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accgggggtac ggcgggacgt ggctacactt 1380

caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac 1500
gagcttaggc ttcccgctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggccctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctacca agctcaagaa cacctacgtg gacccccctc caagcctcgt ccacccgagg 1680
acggggccgcc tccacaccgg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgaccca acctgcagaa catccccgtc cgcacccccct tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
cgcgctctcg cccacctctc cggggacgaa aacctgatca gggctttcca ggaggggaag 1920
gacatccaca ccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
ctgatgcgcc gggcgggcaa gacggtgaac ttcggcgctc tctacggcat gtccgccccat 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgcact caacgcccgg 2220
gtgaagagcg tcaggaggc cgcgagcgc atggccttca acatgccgt ccagggcacc 2280
gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gcccgcagtc tcctccaggt cgccaacgag ctctccttgg agggccccca agcgcggggc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatccccct cgccgtgcc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggta ccaccaccac 2520
caccac 2526

<210> 212

<211> 842

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 212

Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
1 5 10 15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
20 25 30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
35 40 45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
50 55 60

Val	Phe	Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	65	70	75	80
Tyr	Glu	Ala	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	85	90	95	
Arg	Gln	Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr	100	105	110	
Arg	Leu	Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Thr	Leu	115	120	125	
Ala	Lys	Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	130	135	140	
Asp	Arg	Asp	Leu	Tyr	Gln	Leu	Val	Ser	Asp	Arg	Val	Ala	Val	Leu	His	145	150	155	160
Pro	Glu	Gly	His	Leu	Ile	Thr	Pro	Glu	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	165	170	175	
Leu	Arg	Pro	Glu	Gln	Trp	Val	Asp	Phe	Arg	Ala	Leu	Val	Gly	Asp	Pro	180	185	190	
Ser	Asp	Asn	Leu	Arg	Gly	Val	Arg	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Leu	195	200	205	
Lys	Leu	Leu	Lys	Glu	Trp	Gly	Ser	Leu	Glu	Asn	Leu	Leu	Lys	Asn	Leu	210	215	220	
Asp	Arg	Val	Lys	Pro	Glu	Asn	Val	Arg	Glu	Lys	Ile	Lys	Ala	His	Leu	225	230	235	240
Glu	Asp	Leu	Arg	Leu	Ser	Leu	Glu	Leu	Ser	Arg	Val	Arg	Thr	Asp	Leu	245	250	255	
Pro	Leu	Glu	Val	Asp	Leu	Ala	Gln	Gly	Arg	Glu	Pro	Asp	Arg	Glu	Gly	260	265	270	
Leu	Arg	Ala	Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	275	280	285	
Phe	Gly	Leu	Leu	Glu	Ala	Pro	Ala	Pro	Leu	Glu	Glu	Ala	Pro	Trp	Pro	290	295	300	
Pro	Pro	Glu	Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Pro	Glu	Pro	305	310	315	320
Met	Trp	Ala	Glu	Leu	Lys	Ala	Leu	Ala	Ala	Cys	Arg	Gly	Gly	Arg	Val	325	330	335	
His	Arg	Ala	Ala	Asp	Pro	Leu	Ala	Gly	Leu	Lys	Asp	Leu	Lys	Glu	Val	340	345	350	
Arg	Gly	Leu	Leu	Ala	Lys	Asp	Leu	Ala	Val	Leu	Ala	Ser	Arg	Glu	Gly	355	360	365	
Leu	Asp	Leu	Val	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	370	375	380	
Asp	Pro	Ser	Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	385	390	395	400

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
 785 790 795 800
 Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
 805 810 815
 Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
 820 825 830
 Ser Ala Lys Gly His His His His His His
 835 840

<210> 213
 <211> 2526
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 213
 atgaattccg aggcgatgct tccgctcttt gaaccctaaag gccgggtcct cctggtggac 60
 ggccaccacc tggcctaccg caccttcttc gcctgaagg gcctcaccac gagccggggc 120
 gaaccgggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg cgtctctcgt ggtctttgac gccaaggccc cctccttcg ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccc acccccgagg acttcccccg gcagctcgcc 300
 ctcatcaagg agctggtgga cctcctgggg tttaccgcc tcgaggtccc cggctacgag 360
 gcggacgacg ttctcgccac cctggccaag aaggcgaaa aggagggtta cgaggtgcgc 420
 atcctcaccg ccgaccgga cctctaccaa ctctctccg accgcgtcgc cgtcctccac 480
 cccgaggggc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttccgcgc cctcgtgggg gaccctccg acaacctccc cggggtcaag 600
 ggcatcgggg agtataccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
 gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
 gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg 840
 gagttcggca gcctcctcca cgagttcggc ctctggagg ccccgcccc cctggaggag 900
 gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctccg ccccgagccc 960


```

atgtgggcgg agcttaaagc cctggccgcc tgcaggggcg gccgcgtgca ccgggcagca 1020
gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
gccgtcttgg cctcgaggga ggggctagac ctctgtccccg gggacgacct catgctcctc 1140
gcctacctcc tggaccttcc gaacaccacc cccgaggggg tggcgcgggc ctacgggggg 1200
gagtggacgg aggacgccgc ccaccggggc ctctcctcgg agaggctcca tcggaacctc 1260
cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
caggcccttt ccttgagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
gagcttaggc ttcccgctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
gcggtgctgg aggcctacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
gagctcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccaccgagg 1680
acgggcccgc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
tccgacccca acctgcagaa catcccgcgc cgcaccccct tgggccagag gatccgccgg 1800
gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc 1860
cgcgtcctcg ccacctctc cggggacgaa aacctgatca gggctctcca ggagggggaa 1920
gacatccaca ccagaccgc aagctggatg ttcggcgctc ccccgaggc cgtggacccc 1980
ctgatgcgcc gggcgggcaa gacggtgaac ttcggcgctc tctacggcat gtccgcccac 2040
aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
ttccaaagct tccccagggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccaccc caacgcccgg 2220
gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340
gccgcgatgc tcctccaggt cgccaacgag ctctcctcgg agggccccc agcgcgggcc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaagggtca ccaccaccac 2520
caccac

```

2526

```

<210> 214
<211> 842
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

```

<400> 214

Met	Asn	Ser	Glu	Ala	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val
1				5					10					15	
Leu	Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	Phe	Ala	Leu
			20					25					30		
Lys	Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly
		35					40					45			
Phe	Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Tyr	Lys	Ala
	50					55					60				
Val	Phe	Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala
65					70					75					80
Tyr	Glu	Ala	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro
				85					90					95	
Arg	Gln	Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr
			100					105					110		
Arg	Leu	Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Thr	Leu
			115				120					125			
Ala	Lys	Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala
	130					135					140				
Asp	Arg	Asp	Leu	Tyr	Gln	Leu	Val	Ser	Asp	Arg	Val	Ala	Val	Leu	His
145					150				155						160
Pro	Glu	Gly	His	Leu	Ile	Thr	Pro	Glu	Trp	Leu	Trp	Glu	Lys	Tyr	Gly
				165					170					175	
Leu	Arg	Pro	Glu	Gln	Trp	Val	Asp	Phe	Arg	Ala	Leu	Val	Gly	Asp	Pro
			180					185					190		
Ser	Asp	Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Tyr	Thr	Ala	Leu
		195					200					205			
Lys	Leu	Leu	Lys	Glu	Trp	Gly	Ser	Leu	Glu	Asn	Leu	Leu	Lys	Asn	Leu
	210					215					220				
Asp	Arg	Val	Lys	Pro	Glu	Asn	Val	Arg	Glu	Lys	Ile	Lys	Ala	His	Leu
225					230					235					240
Glu	Asp	Leu	Arg	Leu	Ser	Leu	Glu	Leu	Ser	Arg	Val	Arg	Thr	Asp	Leu
				245					250					255	
Pro	Leu	Glu	Val	Asp	Leu	Ala	Gln	Gly	Arg	Glu	Pro	Asp	Arg	Glu	Gly
			260					265					270		
Leu	Arg	Ala	Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu
		275					280					285			
Phe	Gly	Leu	Leu	Glu	Ala	Pro	Ala	Pro	Leu	Glu	Glu	Ala	Pro	Trp	Pro
	290					295					300				
Pro	Pro	Glu	Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Pro	Glu	Pro
305					310					315					320
Met	Trp	Ala	Glu	Leu	Lys	Ala	Leu	Ala	Ala	Cys	Arg	Gly	Gly	Arg	Val

325										330					335				
His	Arg	Ala	Ala	Asp	Pro	Leu	Ala	Gly	Leu	Lys	Asp	Leu	Lys	Glu	Val				
			340					345					350						
Arg	Gly	Leu	Leu	Ala	Lys	Asp	Leu	Ala	Val	Leu	Ala	Ser	Arg	Glu	Gly				
		355					360					365							
Leu	Asp	Leu	Val	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu				
	370					375					380								
Asp	Pro	Ser	Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly				
385					390					395					400				
Glu	Trp	Thr	Glu	Asp	Ala	Ala	His	Arg	Ala	Leu	Leu	Ser	Glu	Arg	Leu				
				405					410					415					
His	Arg	Asn	Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Lys	Leu	Leu	Trp				
			420					425					430						
Leu	Tyr	His	Glu	Val	Glu	Lys	Pro	Leu	Ser	Arg	Val	Leu	Ala	His	Met				
		435					440					445							
Glu	Ala	Thr	Gly	Val	Arg	Arg	Asp	Val	Ala	Tyr	Leu	Gln	Ala	Leu	Ser				
		450					455				460								
Leu	Glu	Leu	Ala	Glu	Glu	Ile	Arg	Arg	Leu	Glu	Glu	Glu	Val	Phe	Arg				
465					470					475					480				
Leu	Ala	Gly	His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg				
				485				490					495						
Val	Leu	Phe	Asp	Glu	Leu	Arg	Leu	Pro	Ala	Leu	Gly	Lys	Thr	Gln	Lys				
			500					505					510						
Thr	Gly	Lys	Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu				
		515					520					525							
Ala	His	Pro	Ile	Val	Glu	Lys	Ile	Leu	Gln	His	Arg	Glu	Leu	Thr	Lys				
	530					535					540								
Leu	Lys	Asn	Thr	Tyr	Val	Asp	Pro	Leu	Pro	Ser	Leu	Val	His	Pro	Arg				
545					550					555					560				
Thr	Gly	Arg	Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly				
			565					570						575					
Arg	Leu	Ser	Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr				
			580					585					590						
Pro	Leu	Gly	Gln	Arg	Ile	Arg	Arg	Ala	Phe	Val	Ala	Glu	Ala	Gly	Trp				
		595					600					605							
Ala	Leu	Val	Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala				
	610					615					620								
His	Leu	Ser	Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Lys				
625					630					635					640				
Asp	Ile	His	Thr	Gln	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Pro	Glu				
				645					650					655					

Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
660 665 670

Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
675 680 685

Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
690 695 700

Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
705 710 715 720

Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
725 730 735

Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
740 745 750

Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
755 760 765

Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
770 775 780

Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
785 790 795 800

Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
805 810 815

Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
820 825 830

Ser Ala Lys Gly His His His His His His
835 840

<210> 215
<211> 2526
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 215
atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
gggtacaagg ccgtcttcgt ggtctttgac gccaaggccc cctccttcg ccacgaggcc 240
tacgaggcct acaaggcggg gagggcccg acccccgagg acttcccccg gcagctcgcc 300
ctcatcaagg agctggtgga cctcctgggg ttaccgccg tcgaggtccc cggctacgag 360
gcgacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc 420
atcctcaccg ccgaccgga cctctaccaa ctgctctccg accgcgtcgc cgtcctccac 480

cccgaggggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag 540
 cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cgggggtcaag 600
 ggcatcaggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg 720
 gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggagggtg 780
 gacctcgccc aggggaggga gcccagaccg gaggggctta gggccttcct ggagaggctg 840
 gagttcggca gcctcctcca cgagttcggc ctcttgagg ccccgcccc cctggaggag 900
 gccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc 960
 atgtgggagg agcttaaagc cctggccgcc tgcaggggag gccgcgtgca ccgggcagca 1020
 gaccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttgg cctcgaggga ggggctagac ctctgcccc gggacgaccc catgctcctc 1140
 gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcgcgcg ctacgggggg 1200
 gagtggacgg aggacgccgc ccaccgggcc ctctctcgg agaggctcca tcggaacctc 1260
 cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
 ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggcctacctt 1380
 caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
 ttggcggggc accccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac 1500
 gagcttaggc ttccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc 1560
 gcggtgctgg aggccttacg ggaggccac cccatcgtgg agaagatcct ccagcaccgg 1620
 gagtcacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg 1680
 acgggcccgc tccacaccgc cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
 tccgacccca acctgcagaa catccccgtc cgcacccctc tgggcccagag gatccgccgg 1800
 gccttcgtgg ccgaggcggg ttgggcttg gtggccttg actatagcca gatagagctc 1860
 cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggcttcca ggaggggaag 1920
 gacatccaca ccagaccgc aagctggatg ttcggcgtcc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccac 2040
 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac 2100
 ttccaaagct tcccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccacct caacgcccgg 2220
 gtgaagagcg tcaggagggc cgcggagcgc atggccttca acatgccgt ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340

```

gcccccatgc tcctccaggt cgccaacgag ctctctctgg agggccccca agcgcggggcc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaaggggtca ccaccaccac 2520
caccac 2526

```

```

<210> 216
<211> 842
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic

```

```

<400> 216
Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
 1          5          10          15

Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
 20          25          30

Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
 35          40          45

Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
 50          55          60

Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
 65          70          75          80

Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
 85          90          95

Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
100          105          110

Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
115          120          125

Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
130          135          140

Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
145          150          155          160

Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
165          170          175

Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
180          185          190

Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Arg Glu Lys Thr Ala Leu
195          200          205

Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
210          215          220

Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
225          230          235          240

```

Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
 245 250 255
 Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
 325 330 335
 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
 355 360 365
 Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
 420 425 430
 Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
 435 440 445
 Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
 450 455 460
 Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
 530 535 540
 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575

caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cggggccgggc cccacgcgcg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggttcacg cgctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc tcccgggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgocg accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg cttcttgga agccccaagg ccctggagga ggccccctgg 900
ccccgcgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct ctttggtt taccgggagg tggagaggcc ctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccacccag gacgggccgc 1680
ctccacaccc gcttcaacca gacggccaag gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcatc 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggagggggc ggacatccac 1920

acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtaggaccc cctgatgcgc 1980
 cgggaggcca agaccatcaa cttcgggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccattccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaggagc accaccacca ccaccac 2517

<210> 218

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 218

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
	145				150					155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
				165					170					175	

Arg Ser Ser Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 219
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 219
cgggacctcg aggcgcgtga accccaggag gtccac 36

<210> 220
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 220
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggcttacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cgagacggag gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcttttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200

gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtccctggccc atatggaggc cacgggggtg cgccctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
 cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcg cgccgtcctg 1560
 gaggcctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacctcttg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc cccggggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttcccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaggagc accaccacca ccaccac 2517

<210> 221

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 221

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Glu Thr Glu Glu Asp Phe Pro Arg Gln
85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
145 150 155 160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
180 185 190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
195 200 205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
210 215 220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
225 230 235 240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
245 250 255

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
260 265 270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
275 280 285

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
290 295 300

Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
305 310 315 320

Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
325 330 335

Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
340 345 350

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
355 360 365

Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
370 375 380

Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Lys Leu Lys Lys Thr Lys Lys Thr Gly Lys
 500 505 510
 Arg Ser Ser Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720

Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
725 730 735

Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
740 745 750

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Gly Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 222
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 222
ccggggaaag tcctcctccg tctcggcccg gccgcctt 39

<210> 223
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 223
gtcggactcg tcaccggtca gggc 24

<210> 224
<211> 75
<212> DNA
<213> Artificial Sequence

<220>
<221> modified_base
<222> (28)..(60)
<223> The bases in these positions within this primer
are 91% of the base shown and 3% each of the other
3 nucleotides.

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 224

ctgaccggtg acgagtcgga caaccttccc ggggtcaagg gcatcgggga gaggacggcg 60
aggaagcttc tggag 75

<210> 225

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 225

atgaattcgg ggatgctgcc cctctttgag cccaagggcc ggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggg cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgctcagagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caatggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caagggtgcg accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggccc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgagc tggcctatct cagggccttg 1380

tccttgaggg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
 cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccccttg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcatc 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggagc ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgccggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgccccagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttccagg tcgccaacga gctggctctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggctgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 226
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 226
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140
 Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Asn Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 227

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 227

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cgggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgcgaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cggggccgggc cccacgccc gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgctcagagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcctcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcac 600
ggggagaaga cgcagaggaa gcttctggag gagggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020

aaagccctca gggacctgaa ggaggcgcg gggtcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctccccgcc ggcgacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccttttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct 1320
 gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
 tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
 cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccttctg ccggacctca tccaccccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccc ggccttcatc 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gcccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgct ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcccggca agaccatcaa cttcggggtc ctctacggca tgcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctccttcagg tcgccaacga gctggtcctc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 228

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 228

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	

Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Lys Leu Lys Lys Thr Lys Lys Thr Gly Lys
 500 505 510
 Arg Ser Ser Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685

Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His
 835

<210> 229

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 229

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cggtctacgg cttoGCCaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacGCCaag gccccctcct tccgccacga ggcctacggg 240
 gggTacaagg cgggCCgggc cccacGCCg gaggactttc cccggcaact cgcctcatc 300
 aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
 gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
 accGCCgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
 gggTacctca tccccCGgc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
 gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatg 600
 ggggagaaga cggggaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660

aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgccc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgctcc accgggcccc cgagccttat 1020
 aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
 gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
 gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
 cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
 gtcctggccc atatggaggc cacgggggtg cgctggacg tggcctatct cagggccttg 1380
 tccttgagg tggccgagga gatcgccgc ctgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
 cttcccaagt tgaagaagac gaagaagacc ggtaagcgt ccagcagcgc cgccgtcctg 1560
 gaggcctcc gcgaggccca cccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
 gccacactct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggat catggagggg gtgtatcccc tggccgtgcc cctggagggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 230
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 230

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
			35				40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85					90					95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
	145				150					155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
			165						170					175	
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp
			180					185					190		
Asn	Leu	Pro	Gly	Val	Lys	Gly	Met	Gly	Glu	Lys	Thr	Gly	Arg	Lys	Leu
		195					200					205			
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg
	210					215					220				
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu
	225				230					235					240
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu
			245						250				255		
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala
			260					265					270		
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu
		275					280					285			

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Lys Leu Lys Lys Thr Lys Lys Thr Gly Lys
 500 505 510
 Arg Ser Ser Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser
 530 535 540
 Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 231

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 231

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
 caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
 ccggtgcagg cggtctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
 gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
 ggggtacaagg cgggccgggc cccacgccg gaggactttc cccggcaact cgcctcatc 300

aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggcctttg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaata cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgcg cccgacctgc ccctggagggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagagggt tgagtctggc 840
agcctcctcc acgagttcgg ccttctggaa agcccaagg ccctggaggga gggccctgg 900
ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380
tccctggagg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
cttcccaagt tgaagaagac gaagaagacc ggtaagcgt ccagcagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgacccttg ccggacctca tccacccag gacgggcccgc 1680
ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct cagggtgctg 1860
gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggcg ggacatccac 1920
acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca ccgcctctcc 2040
caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160

gtggagaccc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tgcacaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
 gcccggtgg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 232

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 232

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5					10					15	
Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
			20					25					30		
Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			
Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
	65				70					75					80
Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
				85				90						95	
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr	Arg	Leu
			100					105					110		
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys
		115					120					125			
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys
	130					135					140				
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu
	145				150					155					160
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg
				165					170					175	
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp
			180					185					190		
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Asn	Thr	Ala	Arg	Lys	Leu
		195					200						205		
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg

210					215					220					
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu
225					230					235					240
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu
				245					250					255	
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala
			260					265					270		
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu
		275					280					285			
Leu	Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu
	290					295					300				
Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala
305					310					315					320
Asp	Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala
				325					330					335	
Pro	Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu
			340					345					350		
Leu	Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu
		355					360					365			
Pro	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser
		370				375					380				
Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr
385					390					395					400
Glu	Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn
				405					410					415	
Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg
			420				425					430			
Glu	Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr
		435					440					445			
Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val
	450					455					460				
Ala	Glu	Glu	Ile	Ala	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly
465					470					475					480
His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe
				485				490						495	
Asp	Glu	Leu	Arg	Leu	Pro	Lys	Leu	Lys	Lys	Thr	Lys	Lys	Thr	Gly	Lys
			500					505					510		
Arg	Ser	Ser	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro
		515				520					525				
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser
	530					535					540				

Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620
 Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His
 625 630 635 640
 Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
 785 790 795 800
 Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Glu His His His His His His
 835

<210> 233
 <211> 2517
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 233

atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccggc cccacgccg gaggactttc cccggcaact cgccctcatc 300
aaggagctgg tggacctcct ggggttcacg cgcctcgagg tccggggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccacgtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaagc cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccatat ggacgatctg 720
aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
cccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccgccc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccgagggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccctttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cttttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcttgacg tggcctatct cagggccttg 1380
tccttgaggg tggccgagga gatcgccgc ctcgaggccg aggtcttccg cctggccggc 1440
cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
cttccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
aagctgaaga gcacctacat tgaccccttg ccggacctca tccacccag gacgggccgc 1680
ctccacacc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740

aacctccaga acatccccgt ccgcacccccg cttgggcaga ggatccgccc ggcccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gcccacctct ccggcgacga gaacctgac cggttcttcc aggaggggccc ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcgggggtc ctctacggca tgcgggcca ccgcctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggcctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagaccc tcttcggccc cgcgcgtac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcccggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgcgcgcgac 2280
 ctcatgaagc tggctatggt gaagctcttc cccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggtcttc gaggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 234
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 234
 Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1 5 10 15
 Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
 20 25 30
 Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
 35 40 45
 Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
 50 55 60
 Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
 65 70 75 80
 Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
 85 90 95
 Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
 100 105 110
 Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
 115 120 125
 Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
 130 135 140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
 145 150 155 160
 Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
 165 170 175
 Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
 180 185 190
 Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Pro Ala Arg Lys Leu
 195 200 205
 Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
 225 230 235 240
 Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
 245 250 255
 Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
 260 265 270
 Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
 275 280 285
 Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 295 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu
 355 360 365
 Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg
 420 425 430
 Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val
 450 455 460
 Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly
 465 470 475 480

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 235
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 235
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggcccgggc ccccacgccg gaggactttc cccggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggttcacg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatcctc 420
accgccgaca aagaccttta ccagtcctt tccgaccgca tccagtcct ccaccccgag 480
gggtacctca tcaccccggc ctggctttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggt caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660
aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggcccacat ggacgatctg 720
aagctctcct gggacctggc caaggtgcmc accgacctgc ccctggaggt ggacttcgcc 780
aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagagggt tgagtctggc 840
agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
ccccgcggg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtgggcc 960
gatcttctgg ccctggccgc cgccaggggc ggccgcgtcc accgggcccc cgagccttat 1020
aaagccctca gggacctgaa ggaggcgcgg gggcttctcg ccaaagacct gagcgttctg 1080
gccctgaggg aaggccttgg cctcccggcc ggcgacgacc ccatgctcct cgcctacctc 1140
ctggaccctt cgaacaccac ccccaggggg gtggcccggc gctacggcgg ggagtggacg 1200
gaggaggcgg gggagcgggc cgccttttcc gagaggctct tcgccaacct gcttaagagg 1260
cttgaggggg aggagaggct cctttggctt taccgggagg tggagaggcc cctttccgct 1320
gtcctggccc atatggaggc cacgggggtg cgcctggacg tggcctatct cagggccttg 1380

tccctggagg tggccgagga gatcgcccg ctcgaggccg aggtcttccg cctggccggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tcctctttga cgagctcagg 1500
 cttcccaagt tgaagaagac gaagaagacc ggtaagcgct ccagcagcgc cgccgtcctg 1560
 gaggccctcc gcgaggccca ccccatcgtg gagaagatcc tgcagtaccg ggagctcacc 1620
 aagctgaaga gcacctacat tgaccccttg ccggacctca tccacccag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggca ggctaagtag ctccgatccc 1740
 aacctccaga acatccccgt ccgcaccccg cttgggcaga ggatccgccg ggccttcac 1800
 gccgaggagg ggtggctatt ggtggccctg gactatagcc agatagagct caggggtgctg 1860
 gccacctct ccggcgacga gaacctgatc cgggtcttcc aggaggggagc ggacatccac 1920
 acggagaccg ccagctggat gttcggcgtc ccccgaggag ccgtggaccc cctgatgcgc 1980
 cgggcggcca agaccatcaa cttcggggtc ctctacggca tgtcggccca cgccctctcc 2040
 caggagctag ccatccctta cgaggaggcc caggccttca ttgagcgcta ctttcagagc 2100
 ttccccaagg tgcgggctg gattgagaag accctggagg agggcaggag gcgggggtac 2160
 gtggagacc tcttcggccg ccgccgctac gtgccagacc tagaggcccg ggtgaagagc 2220
 gtgcgggagg cggccgagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tggctatggt gaagctcttc ccaggctgg aggaaatggg ggccaggatg 2340
 ctcttcagg tcgccaacga gctggtcctc gagggcccaa aagagagggc ggaggccgtg 2400
 gcccggtg ccaaggaggt catggagggg gtgtatcccc tggccgtgcc cctggaggtg 2460
 gaggtgggga taggggagga ctggctctcc gccaaaggagc accaccacca ccaccac 2517

<210> 236

<211> 839

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 236

Met	Asn	Ser	Gly	Met	Leu	Pro	Leu	Phe	Glu	Pro	Lys	Gly	Arg	Val	Leu
1				5				10						15	

Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	His	Ala	Leu	Lys
		20						25					30		

Gly	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe
		35					40					45			

Ala	Lys	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Asp	Ala	Val	Ile
	50					55					60				

Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Gly
65					70					75				80	

Gly	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln	
				85					90					95		
Leu	Ala	Leu	Ile	Lys	Glu	Leu	Val	Asp	Leu	Leu	Gly	Phe	Thr	Arg	Leu	
			100					105					110			
Glu	Val	Pro	Gly	Tyr	Glu	Ala	Asp	Asp	Val	Leu	Ala	Ser	Leu	Ala	Lys	
		115					120					125				
Lys	Ala	Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Lys	
	130					135					140					
Asp	Leu	Tyr	Gln	Leu	Leu	Ser	Asp	Arg	Ile	His	Val	Leu	His	Pro	Glu	
145					150					155					160	
Gly	Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	
				165					170					175		
Pro	Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	
			180					185					190			
Asn	Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu	
		195					200					205				
Leu	Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg	
	210					215					220					
Leu	Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu	
225					230					235					240	
Lys	Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	
				245					250					255		
Val	Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	
			260					265					270			
Phe	Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	
		275					280					285				
Leu	Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	
	290					295					300					
Gly	Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	
305					310					315					320	
Asp	Leu	Leu	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	
				325					330					335		
Pro	Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu	
			340					345					350			
Leu	Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	
		355					360					365				
Pro	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	
		370				375					380					
Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	
385					390					395					400	
Glu	Glu	Ala	Gly	Glu	Arg	Ala	Ala	Leu	Ser	Glu	Arg	Leu	Phe	Ala	Asn	
				405					410					415		

425
 430
 435
 440
 445
 450
 455
 460
 465
 470
 475
 480
 485
 490
 495
 500
 505
 510
 515
 520
 525
 530
 535
 540
 545
 550
 555
 560
 565
 570
 575
 580
 585
 590
 595
 600
 605
 610
 615
 620
 625
 630
 635
 640
 645
 650
 655
 660
 665
 670
 675
 680
 685
 690
 695
 700
 705
 710
 715
 720
 725
 730
 735
 740
 745
 750

Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Arg		
			420					425					430				
Glu	Val	Glu	Arg	Pro	Leu	Ser	Ala	Val	Leu	Ala	His	Met	Glu	Ala	Thr		
		435					440					445					
Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Arg	Ala	Leu	Ser	Leu	Glu	Val		
	450					455					460						
Ala	Glu	Glu	Ile	Ala	Arg	Leu	Glu	Ala	Glu	Val	Phe	Arg	Leu	Ala	Gly		
465					470					475					480		
His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu	Phe		
			485					490						495			
Asp	Glu	Leu	Arg	Leu	Pro	Lys	Leu	Lys	Lys	Thr	Lys	Lys	Thr	Gly	Lys		
			500					505					510				
Arg	Ser	Ser	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro		
		515					520					525					
Ile	Val	Glu	Lys	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Thr	Lys	Leu	Lys	Ser		
	530					535					540						
Thr	Tyr	Ile	Asp	Pro	Leu	Pro	Asp	Leu	Ile	His	Pro	Arg	Thr	Gly	Arg		
545					550					555					560		
Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser		
				565					570					575			
Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly		
			580					585					590				
Gln	Arg	Ile	Arg	Arg	Ala	Phe	Ile	Ala	Glu	Glu	Gly	Trp	Leu	Leu	Val		
		595					600					605					
Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser		
		610				615					620						
Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Arg	Asp	Ile	His		
625					630					635				640			
Thr	Glu	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Arg	Glu	Ala	Val	Asp		
				645					650					655			
Pro	Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr		
			660					665					670				
Gly	Met	Ser	Ala	His	Ala	Leu	Ser	Gln	Glu	Leu	Ala	Ile	Pro	Tyr	Glu		
		675					680					685					
Glu	Ala	Gln	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Phe	Pro	Lys	Val		
	690					695					700						
Arg	Ala	Trp	Ile	Glu	Lys	Thr	Leu	Glu	Glu	Gly	Arg	Arg	Arg	Gly	Tyr		
705					710					715					720		
Val	Glu	Thr	Leu	Phe	Gly	Arg	Arg	Arg	Tyr	Val	Pro	Asp	Leu	Glu	Ala		
				725					730					735			
Arg	Val	Lys	Ser	Val	Arg	Glu	Ala	Ala	Glu	Arg	Met	Ala	Phe	Asn	Met		
			740					745					750				

Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
755 760 765

Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val
770 775 780

Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val
785 790 795 800

Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val
805 810 815

Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys
820 825 830

Glu His His His His His His
835

<210> 237
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 237
tagctcctgg gagagggcgt gggccgacat gcc 33

<210> 238
<211> 2517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 238
atgaattcgg ggatgctgcc cctctttgag cccaagggcc gggtcctcct ggtggacggc 60
caccacctgg cctaccgcac cttccacgcc ctgaagggcc tcaccaccag ccgggggggag 120
ccggtgcagg cggctctacgg cttcgccaag agcctcctca aggccctcaa ggaggacggg 180
gacgcggtga tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacggg 240
gggtacaagg cgggccgggc cccacgccg gaggactttc ccgggcaact cgcctcatc 300
aaggagctgg tggacctcct ggggctggcg cgcctcgagg tcccgggcta cgaggcggac 360
gacgtcctgg ccagcctggc caagaaggcg gaaaaggagg gctacgaggt ccgcatactc 420
accgccgaca aagaccttta ccagctcctt tccgaccgca tccagctcct ccaccccgag 480
gggtacctca tcaccccggc ctggccttgg gaaaagtacg gcctgaggcc cgaccagtgg 540
gccgactacc gggccctgac cggggacgag tccgacaacc ttcccggggg caagggcatc 600
ggggagaaga cggcgaggaa gcttctggag gagtggggga gcctggaagc cctcctcaag 660

aacctggacc ggctgaagcc cgccatccgg gagaagatcc tggccacat ggacgatctg 720
 aagctctcct gggacctggc caaggtgcgc accgacctgc ccctggaggt ggacttcgcc 780
 aaaaggcggg agcccgaccg ggagaggctt agggcctttc tggagaggct tgagtttggc 840
 agcctcctcc acgagttcgg ccttctggaa agccccaagg ccctggagga ggccccctgg 900
 ccccgccgg aaggggcctt cgtgggcttt gtgctttccc gcaaggagcc catgtggggc 960
 gatcttctgg ccctggccgc cgccaggggc ggccgcgtgc accgggcagc agacccttg 1020
 gcggggctaa aggacctcaa ggaggtcgg ggccctctcg ccaaggacct cgccgtcttg 1080
 gcctcgaggg aggggctaga cctcgtgccc ggggacgacc ccatgctcct cgcctacctc 1140
 ctggaccctt cgaacaccac ccccgagggg gtggcgcggc gctacggggg ggagtggacg 1200
 gaggacgccc cccaccgggc cctcctctcg gagaggctcc atcggaacct ccttaagcgc 1260
 ctcgaggggg aggagaagct cctttggctc taccacgagg tggaaaagcc cctctcccgg 1320
 gtcctggccc atatggaggc caccggggta cggcgggacg tggcctacct tcaggccctt 1380
 tccctggagc ttgctggagga gatccgccgc ctcgaggagg aggtcttccg cttggcgggc 1440
 cacccttca acctcaactc ccgggaccag ctggaaaggg tgctctttga cgagcttagg 1500
 cttccgcct tgaagaagac gaagaagaca ggcaagcgct ccaccagcgc cgcggtgctg 1560
 gaggccctac gggaggccca ccccatcgtg gagaagatcc tccagcaccg ggagctcacc 1620
 aagctcaaga acacctacgt ggacccccctc ccaagcctcg tccacccgag gacgggcccgc 1680
 ctccacaccc gcttcaacca gacggccacg gccacgggga ggcttagtag ctccgacccc 1740
 aacctgcaga acatccccgt ccgcaccccc ttgggccaga ggatccgccg ggccttcgtg 1800
 gccgaggcgg gttgggcgtt ggtggccctg gactatagcc agatagagct ccgcgtcctc 1860
 gccacctct ccggggacga aaacctgatc agggcttcc aggaggggaa ggacatccac 1920
 acccagaccg caagctggat gttcggcgtc ccccgaggagg ccgtggaccc cctgatgcgc 1980
 cgggcggcca agacggtgaa cttcggcgtc ctctacggca tgtccgcca taggctctcc 2040
 caggagcttg ccatccccta cgaggaggcg gtggccttta tagagcgcta cttccaaagc 2100
 tcccccaagg tgcgggcctg gatagaaaag accctggagg aggggaggaa gcggggctac 2160
 gtggaaaccc tcttcggaag aaggcgctac gtgcccagcc tcaacgcccc ggtgaagagc 2220
 gtcaggagg ccgcggagcg catggccttc aacatgcccg tccagggcac cgccgccgac 2280
 ctcatgaagc tcgccatggt gaagctcttc cccgcctcc gggagatggg ggcccgcatg 2340
 ctctccagg tcgccaacga gctcctcctg gagggcccc aagcgcgggc cgaggaggtg 2400
 gcggcttttg ccaaggaggc catggagaag gcctatcccc tcgccgtgcc cctggaggtg 2460
 gaggtgggga tgggggagga ctggccttcc gccaaagggtc accaccacca ccaccac 2517

<210> 239
 <211> 839
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 239

```

Met Asn Ser Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu
 1           5           10           15

Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys
          20           25           30

Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe
          35           40           45

Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile
          50           55           60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly
          65           70           75           80

Gly Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
          85           90           95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu
          100          105          110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys
          115          120          125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys
          130          135          140

Asp Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu
          145          150          155          160

Gly Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg
          165          170          175

Pro Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp
          180          185          190

Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu
          195          200          205

Leu Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg
          210          215          220

Leu Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu
          225          230          235          240

Lys Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu
          245          250          255

Val Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala
          260          265          270

Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu
          275          280          285
  
```

Leu Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu
 290 325 300
 Gly Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala
 305 310 315 320
 Asp Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala
 325 330 335
 Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly Leu
 340 345 350
 Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp Leu
 355 360 365
 Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser
 370 375 380
 Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr
 385 390 395 400
 Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg Asn
 405 410 415
 Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr His
 420 425 430
 Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr
 435 440 445
 Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Leu
 450 455 460
 Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala Gly
 465 470 475 480
 His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe
 485 490 495
 Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys Thr Gly Lys
 500 505 510
 Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro
 515 520 525
 Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys Asn
 530 535 540
 Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly Arg
 545 550 555 560
 Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser
 565 570 575
 Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly
 580 585 590
 Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu Val
 595 600 605
 Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
 610 615 620

Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile His
 625 630 635 640
 Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp
 645 650 655
 Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr
 660 665 670
 Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu
 675 680 685
 Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val
 690 695 700
 Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr
 705 710 715 720
 Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala
 725 730 735
 Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met
 740 745 750
 Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys
 755 760 765
 Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln Val
 770 775 780
 Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu Val
 785 790 795 800
 Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala Val
 805 810 815
 Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys
 820 825 830
 Gly His His His His His His
 835

<210> 240

<211> 2526

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 240

atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac 60
 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc 120
 gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac 180
 ggggtacaagg ccgtcttcgt ggtctttgac gccaaaggccc cctccttcgc ccacgaggcc 240
 tacgaggcct acaaggcggg gagggccccg acccccaggg acttccccgc gcagctcgcc 300

ctcatcaagg agctggtgga cctcctggggg tttacccggc tcgaggtccc cggctacgag 360
gcggaacgacg ttctcgccac cctggccaag aaggcggaag aggaggggta cgaggtgagc 420
atcctcaccg ccgaccgga cctctaccaa ctctgtctccg accggtgagc cgtcctccac 480
cccgaggggc acctcatcac ccgggagtg ctttgggaga agtacggcct caggccggag 540
cagtgggtgg acttccgagc cctcgtgggg gacccctccg acaacctccc cgggggtcaag 600
ggcatcgggg agaagaccg cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc 660
ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggccacctg 720
gaagacctca ggctctcctt ggagctctcc cgggtgagca ccgacctccc cctggaggtg 780
gacctcgccc aggggagggg gcccgaccg gaggggctta gggccttccg ggagaggtg 840
gagttcgga gcctcctcca cgagttcggc ctctggagg ccccgcccc cctggaggag 900
gccccctggc ccccgccgga aggggccttc gtgggcttcg tctctcccg ccccgagccc 960
atgtgggagg agcttaaagc cctggccgac tgcaggggag gccggtcca ccgggcccc 1020
gagccttata aagccctcag ggacctgaag gaggcgaggg ggcttctcgc caaagacctg 1080
agcgttctgg ccctgagggg aggccttggc ctcccgccc ggcagacccc catgctcctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcccgagc ctacggcggg 1200
gagtggacgg aggggagggg ggagcggggc gccctttccg agaggctctt cgccaacctg 1260
cttaagaggc ttgaggggga ggagaggtc ctttggcttt accgggaggt ggagaggccc 1320
ctttccgctg tcttgccca tatggaggc acgggggtgc gcctggacgt ggcttatctc 1380
agggccttgt ccctggaggt ggccgaggag atcgcccgcc tcgaggccga ggtcttccgc 1440
ctggccgggc accccttcaa cctcaactcc cgggaccagc tggaaagggt cctctttgac 1500
gagctagggc ttcccgccat caagaagacg caaaagaccg gcaagcgtc caccagcgcc 1560
gccgtcctgg aggcctcctg cgaggccac cccatcgtg agaagatcct gcagtaccg 1620
gagctcacca agctgaagag cacctacatt gacccttgc cggacctcat ccacccag 1680
acgggcccgc tccacaccg cttcaaccag acggccacgg ccacgggag gctaagtagc 1740
tccgatccca acctccagaa catccccgtc cgcacccgc ttgggagag gatccggc 1800
gccttcatcg ccgaggaggg gtggctattg gtggccctg actatagcca gatagagctc 1860
aggggtgctg cccacctctc cggcgacgag aacctgatcc gggcttcca ggaggggagc 1920
gacatccaca cggagaccg cagctggatg ttggcgctc cccgggaggc cgtggacccc 1980
ctgatgcgcc gggcgccaa gaccatcaac ttgggggtcc tctacggcat gtcggccac 2040
cgcctctccc aggagctagc catcccttac gaggaggccc aggccttcat tgagcgctac 2100
tttcagagct tccccagggt gcgggcctg attgagaaga ccctggagga gggcaggag 2160

Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
 210 215 220
 Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
 225 230 235 240
 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
 245 250 255
 Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Gly Gly Arg Val
 325 330 335
 His Arg Ala Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly
 355 360 365
 Leu Gly Leu Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu
 405 410 415
 Phe Ala Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp
 420 425 430
 Leu Tyr Arg Glu Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met
 435 440 445
 Glu Ala Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser
 450 455 460
 Leu Glu Val Ala Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Gly Leu Pro Ala Ile Lys Lys Thr Gln Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys
 530 535 540

Leu Lys Ser Thr Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575
 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp
 595 600 605
 Leu Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg
 625 630 635 640
 Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val Ala Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala
 785 790 795 800
 Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro
 805 810 815
 Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu
 820 825 830
 Ser Ala Lys Glu His His His His His His
 835 840

<210> 242
 <211> 2508
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 242

atggaattca cccactttt tgacctggag gaacccccca agcgggtgct tctggtggac 60
ggccaccacc tggcctaccg caccttctat gccctgagcc tcaccacctc cgggggggag 120
ccggtgcaga tggctctacg cttcgcccgg agcctcctca aggccttgaa ggaggacgga 180
caggcgggtg tctggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
gcctacaagg cgggccgggc cccacccccg gaggacttcc cccgccagct cgccttggtc 300
aagcggctgg tggaccttct gggcctggtc cgcctcgagg ccccggggta cgaggcggac 360
gacgtcctgg gcaccctggc caagaaggcc gaaaggagg ggatggaggt gcgcacctc 420
acgggagacc gggacttctt ccagctcctc tccgagaagg tctcggtcct cctgccggac 480
gggaccctgg tcaccccaaa ggacgtccag gagaagtacg gggtgcccc ggagcgctgg 540
gtggacttcc gcgccctcac gggggaccgc tcggacaaca tccccggggt ggcggggata 600
ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
ctccacctct ccttagacct ggcccgcatc cgcaccgacc tccccctgga ggtggacttt 780
aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840
gagttcggaa gcctcctcca cgagtccggc ctctggggag gggagaagcc cggggaggag 900
gccccctggc ccccgccga aggggccttc gtgggcttcc tcttttccg caaggagccc 960
atgtgggcgg agcttctggc cctggcgggc gcctcgagg gccgggtcca cggggcaaca 1020
agcccggttg aggccctggc cgacctcaag gaggcccggg ggttcctggc caaggacctg 1080
gccgttttgg ccctgcggga gggggtggcc ctggaccca cggacgacct cctcctggtg 1140
gcctacctcc tggaccgggc caacaccac cccgaggggg tggcccggg ctacgggggc 1200
gagttcacgg aggacgcagc ggagagggcc ctctctccg agaggctctt ccagaacctc 1260
tttaaacggc tttccgagaa gctcctctgg ctctaccagg aggtggagcg gccctctcc 1320
cgggtcttgg ccacatgga gggccggggg gtgaggtgg acgtccccct tctggaggcc 1380
ctctcctttg agctggagaa ggagatggag cgctggagg gggaggtctt ccgtttggcc 1440
ggcaccctc tcaacctcaa ctcccgac cagctggaaa ggtcctctt tgacgagctg 1500
ggcctcacc cgggtggccg gacgcagaag acgggcaagc gctccaccgc ccagggggcc 1560
ctggaggccc tccggggggc ccacccatc gtggagctca tcctccagta ccgggagctt 1620
tccaagctca aaagcaccta cctggacccc ctgccccggc tcgtccacc gcggacgggc 1680
cggctccaca cccgcttcaa ccagacggcc acggccacgg gaaggcttcc cagctccgac 1740

```

cccaacctgc agaacatccc cgtgcgacc cccttggggc agcgcacccg caaggccttc 1800
gtggccgagg aggggtggct ccttttggcg gcggactact cccagattga gctccgggtc 1860
ctggcccacc tctcggggga cgagaacctg aagcgggtct tccgggaggg gaaggacatc 1920
cataccgaga ccgccgcctg gatgttcggc ttagaccccg ctctggtgga tccaaagatg 1980
cgccggggcg ccaagacggt caacttcggc gtcctctacg ggatgtccgc ccacagggtc 2040
tcccaggagc tcggcataga ctacaaggag gcggaggcct ttattgagcg ctacttccag 2100
agcttcccca aggtgcgggc ctggatagaa aggaccctgg aggagggccg gacgcggggc 2160
tacgtggaga ccctgttcgg caggaggcgc tatgtgcccg acctggcctc ccgggtccgc 2220
tcggtgcggg aggcggcgga gcggatggcc ttcaacatgc ccgtgcaggg caccgccgcc 2280
gacctgatga agatcgccat ggtcaagctc ttccccaggc taaagcccct gggggcccac 2340
ctcctcctcc aagtgcacaa cgagctggtc ctggaggtgc ccgaggaccg ggccgaggag 2400
gccaaggccc tggtaagga ggtcatggag aacgcctacc ccctggacgt gcccctcgag 2460
gtggaggtgg gcgtgggtcg ggactggctg gaggcgaagc aggattga 2508

```

<210> 243

<211> 835

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 243

```

Met Glu Phe Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
  1             5             10             15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
          20             25             30
Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
          35             40             45
Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
          50             55             60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
          65             70             75             80
Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
          85             90             95
Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg Leu
          100            105            110
Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
          115            120            125
Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
          130            135            140

```

Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
 145 150 155 160
 Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
 165 170 175
 Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
 180 185 190
 Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
 195 200 205
 Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
 225 230 235 240
 Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ser Glu Gly Arg Val
 325 330 335
 His Arg Ala Thr Ser Pro Val Glu Ala Leu Ala Asp Leu Lys Glu Ala
 340 345 350
 Arg Gly Phe Leu Ala Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly
 355 360 365
 Val Ala Leu Asp Pro Thr Asp Asp Pro Leu Leu Val Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ala Asn Thr His Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Phe Thr Glu Asp Ala Ala Glu Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 Phe Gln Asn Leu Phe Lys Arg Leu Ser Glu Lys Leu Leu Trp Leu Tyr
 420 425 430
 Gln Glu Val Glu Arg Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
 435 440 445
 Arg Gly Val Arg Leu Asp Val Pro Leu Leu Glu Ala Leu Ser Phe Glu
 450 455 460
 Leu Glu Lys Glu Met Glu Arg Leu Glu Gly Glu Val Phe Arg Leu Ala
 465 470 475 480

Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
485 490 495

Phe Asp Glu Leu Gly Leu Thr Pro Val Gly Arg Thr Gln Lys Thr Gly
500 505 510

Lys Arg Ser Thr Ala Gln Gly Ala Leu Glu Ala Leu Arg Gly Ala His
515 520 525

Pro Ile Val Glu Leu Ile Leu Gln Tyr Arg Glu Leu Ser Lys Leu Lys
530 535 540

Ser Thr Tyr Leu Asp Pro Leu Pro Arg Leu Val His Pro Arg Thr Gly
545 550 555 560

Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
565 570 575

Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
580 585 590

Gly Gln Arg Ile Arg Lys Ala Phe Val Ala Glu Glu Gly Trp Leu Leu
595 600 605

Leu Ala Ala Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
610 615 620

Ser Gly Asp Glu Asn Leu Lys Arg Val Phe Arg Glu Gly Lys Asp Ile
625 630 635 640

His Thr Glu Thr Ala Ala Trp Met Phe Gly Leu Asp Pro Ala Leu Val
645 650 655

Asp Pro Lys Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
660 665 670

Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Gly Ile Asp Tyr
675 680 685

Lys Glu Ala Glu Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
690 695 700

Val Arg Ala Trp Ile Glu Arg Thr Leu Glu Glu Gly Arg Thr Arg Gly
705 710 715 720

Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Ala
725 730 735

Ser Arg Val Arg Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
740 745 750

Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Ile Ala Met Val
755 760 765

Lys Leu Phe Pro Arg Leu Lys Pro Leu Gly Ala His Leu Leu Leu Gln
770 775 780

Val His Asn Glu Leu Val Leu Glu Val Pro Glu Asp Arg Ala Glu Glu
785 790 795 800

Ala Lys Ala Leu Val Lys Glu Val Met Glu Asn Ala Tyr Pro Leu Asp
805 810 815

Val Pro Leu Glu Val Glu Val Gly Val Gly Arg Asp Trp Leu Glu Ala
820 825 830

Lys Gln Asp
835

<210> 244
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 244
cttccagaac ctctttaaac ggctttccga gaag 34

<210> 245
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 245
cttctcggaa agccgtttaa agaggttctg gaag 34

<210> 246
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 246
ccggtgggcc ggacgcagaa gacgggcaag c 31

<210> 247
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 247
gcttgcccgt cttctgcgtc cggcccaccg g 31

<210> 248
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 248
ctcctccaag tgcacaacga gctggctcctg g 31

<210> 249
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 249
ccaggaccag ctcggttggtg acttgaggag g 31

<210> 250
<211> 2499
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 250
atggaattcc tgcccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
ctggcctacc gtaccttttt tgccctgaag ggccctacca ccagccgcgg ggagccgggtc 120
caggcgggtgt acgggttttg caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
gtgatcgtgg tgtttgacgc caaggcccc tccttcgcc accagacctt cgaggcctac 240
aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300
atggtggacc ttttgggcct ggagcgcctc gaggtgccgg gctttgaagc ggatgacgtc 360
ctggctaccc tggccaagaa ggcgaaaag gaaggctacg aagtgcgcac cctcaccgcg 420
gaccgggacc tttaccagct tctttcgag cgaatctcca tccttcaccc ggagggttac 480
ctgatcacc cggagtggct ttgggagaag tatgggctta agccttccca gtgggtggac 540
taccgggcct tggccgggga cccttcgcac aacatccccg gcgtgaaggg catcggggag 600
aagacggcgg ccaagctgat ccgggagtg ggaagcctg aaaccttct taagcacctg 660
gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag 720
ctatccctgg agctatccc ggtgcacacg gacttgctcc ttcagggtgga cttcgccccg 780
cgccgggagc cggaccggga ggggcttaag gcctttttgg agaggctgga gttcggaagc 840
ctcctccacg agttcggcct gttggaaaag ccggtggcgg cggaggaagc tccctggccg 900
cccccgagg gagccttcgt ggggtacgtt ctttcccgcc ccgagcccat gtgggcggag 960
cttaacgcct tggccgccgc ctgggagggg agggtttacc gggcgaggga tcccttgagg 1020
gccttgccgg ggcttgggga ggtgaggggg cttttggcca aggacctggc ggtgctggcc 1080
ctgaggggaag ggattgccct ggcaccgggc gacgaccca tgctcctcgc ctacctcctg 1140

gatccttcca acaccgcccc cgaaggggta gcccgcgct acggggggga gtggaccgag 1200
gaggcggggg aaagggcgct gctttccgaa aggcctttacg ccgccctcct gaagcggctt 1260
aagggggagg agaggcttct ttggctttac gaggaggtgg aaaagccctt ttcgcgggtc 1320
ctggcccaca tggaggccac ggggggtacgg ttggatgtgg cctacttaaa ggccctttcc 1380
ctggaggtgg aggcgggagat aaggcgcttc gaggaggagg tccaccgcct ggccgggcat 1440
cctttcaacc tgaactcccc ggaccagctg gaaaggggtca tctttgacga gcttgggctt 1500
cccgccatcg gcaagacgca gaagacgggc aagcgctcca ccagcgccgc cgttttggag 1560
gccttgcggg aggcctcatcc catcgtggac cgcctccttc agtaccggga gctttccaag 1620
ctcaaggga cctacatcga tcccttgctt gccctgggtcc accccaagac gaaccgcctc 1680
cacaccggtt tcaaccagac ggccaccgcc acggggaggc ttagcagctc ggatcctaata 1740
ctgcaaaata tccccgtgcg cacccttttg ggccagcgga tccgcggggc cttcgtggcc 1800
gaggaggggt ggaggctggt ggttttggac tacagccaga ttgagctcag ggtcctggcg 1860
cacctttccg gggacgagaa cctaataccg gtcttccagg agggccagga catccacacc 1920
cagacggcca gctggatgtt cggcgtgccc ccagaggccg tggattccct gatgcgccgg 1980
gcgccaaga ccatcaactt cggcgtcctc tacggcatgt ccgccaccg gctttcggga 2040
gagctggcca tcccctacga ggaggcgggt gccttcatcg agcgggtattt ccagagctac 2100
cccaagggtgc gggcctggat tgagaaaacc ctggcggaag gacgggaacg gggctatgtg 2160
gaaaccctct ttggccgccc gcgctacgtg ccgacttggt cttcccggtt gaagagcatc 2220
cgggaggcag cggagcgcat ggccttcaac atgcccgttc aggggaccgc cgcggatttg 2280
atgaaactgg ccatggtgaa gctctttccc aggccttcagg agctgggggc caggatgctt 2340
ttgcaggtgc acaacgaact ggtcctcgag gctcccaagg agcaagcgga ggaagtcgcc 2400
caggaggcca agcggacat ggaggaggtg tggcccctga aggtgccctt ggagggtgaa 2460
gtgggcatcg gggaggactg gctttccgcc aaggcctag 2499

<210> 251

<211> 832

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 251

Met Glu Phe Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1 5 10 15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
20 25 30

Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Ala	Val	Tyr	Gly	Phe	Ala	Lys	
		35					40					45				
Ser	Leu	Leu	Lys	Ala	Leu	Arg	Glu	Asp	Gly	Asp	Val	Val	Ile	Val	Val	
	50					55					60					
Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Gln	Thr	Tyr	Glu	Ala	Tyr	
65					70					75					80	
Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln	Leu	Ala	
				85					90					95		
Leu	Ile	Lys	Glu	Met	Val	Asp	Leu	Leu	Gly	Leu	Glu	Arg	Leu	Glu	Val	
			100					105					110			
Pro	Gly	Phe	Glu	Ala	Asp	Asp	Val	Leu	Ala	Thr	Leu	Ala	Lys	Lys	Ala	
	115						120					125				
Glu	Lys	Glu	Gly	Tyr	Glu	Val	Arg	Ile	Leu	Thr	Ala	Asp	Arg	Asp	Leu	
130						135					140					
Tyr	Gln	Leu	Leu	Ser	Glu	Arg	Ile	Ser	Ile	Leu	His	Pro	Glu	Gly	Tyr	
145					150					155					160	
Leu	Ile	Thr	Pro	Glu	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Lys	Pro	Ser	
				165					170					175		
Gln	Trp	Val	Asp	Tyr	Arg	Ala	Leu	Ala	Gly	Asp	Pro	Ser	Asp	Asn	Ile	
			180					185					190			
Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Ala	Lys	Leu	Ile	Arg	
	195						200					205				
Glu	Trp	Gly	Ser	Leu	Glu	Asn	Leu	Leu	Lys	His	Leu	Glu	Gln	Val	Lys	
210						215					220					
Pro	Ala	Ser	Val	Arg	Glu	Lys	Ile	Leu	Ser	His	Met	Glu	Asp	Leu	Lys	
225					230					235					240	
Leu	Ser	Leu	Glu	Leu	Ser	Arg	Val	His	Thr	Asp	Leu	Leu	Leu	Gln	Val	
			245					250						255		
Asp	Phe	Ala	Arg	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Gly	Leu	Lys	Ala	Phe	
			260					265					270			
Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	Leu	
	275						280					285				
Glu	Ser	Pro	Val	Ala	Ala	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	Gly	
	290					295					300					
Ala	Phe	Val	Gly	Tyr	Val	Leu	Ser	Arg	Pro	Glu	Pro	Met	Trp	Ala	Glu	
305					310					315					320	
Leu	Asn	Ala	Leu	Ala	Ala	Ala	Trp	Glu	Gly	Arg	Val	Tyr	Arg	Ala	Glu	
			325						330					335		
Asp	Pro	Leu	Glu	Ala	Leu	Arg	Gly	Leu	Gly	Glu	Val	Arg	Gly	Leu	Leu	
		340						345					350			
Ala	Lys	Asp	Leu	Ala	Val	Leu	Ala	Leu	Arg	Glu	Gly	Ile	Ala	Leu	Ala	
		355					360					365				

Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	Asn	
	370					375						380				
Thr	Ala	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp	Thr	Glu	
	385				390					395					400	
Glu	Ala	Gly	Glu	Arg	Ala	Leu	Leu	Ser	Glu	Arg	Leu	Tyr	Ala	Ala	Leu	
				405					410					415		
Leu	Lys	Arg	Leu	Lys	Gly	Glu	Glu	Arg	Leu	Leu	Trp	Leu	Tyr	Glu	Glu	
			420					425					430			
Val	Glu	Lys	Pro	Leu	Ser	Arg	Val	Leu	Ala	His	Met	Glu	Ala	Thr	Gly	
		435					440					445				
Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Lys	Ala	Leu	Ser	Leu	Glu	Val	Glu	
	450					455					460					
Ala	Glu	Ile	Arg	Arg	Phe	Glu	Glu	Glu	Val	His	Arg	Leu	Ala	Gly	His	
	465				470					475					480	
Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Ile	Phe	Asp	
				485					490					495		
Glu	Leu	Gly	Leu	Pro	Ala	Ile	Gly	Lys	Thr	Gln	Lys	Thr	Gly	Lys	Arg	
			500					505					510			
Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His	Pro	Ile	
		515					520					525				
Val	Asp	Arg	Ile	Leu	Gln	Tyr	Arg	Glu	Leu	Ser	Lys	Leu	Lys	Gly	Thr	
	530					535					540					
Tyr	Ile	Asp	Pro	Leu	Pro	Ala	Leu	Val	His	Pro	Lys	Thr	Asn	Arg	Leu	
	545				550					555					560	
His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu	Ser	Ser	
				565				570						575		
Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu	Gly	Gln	
			580					585					590			
Arg	Ile	Arg	Arg	Ala	Phe	Val	Ala	Glu	Glu	Gly	Trp	Arg	Leu	Val	Val	
		595					600					605				
Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu	Ser	Gly	
	610					615					620					
Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Gln	Asp	Ile	His	Thr	
	625				630					635					640	
Gln	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Pro	Glu	Ala	Val	Asp	Ser	
				645					650					655		
Leu	Met	Arg	Arg	Ala	Ala	Lys	Thr	Ile	Asn	Phe	Gly	Val	Leu	Tyr	Gly	
			660					665					670			
Met	Ser	Ala	His	Arg	Leu	Ser	Gly	Glu	Leu	Ala	Ile	Pro	Tyr	Glu	Glu	
		675					680					685				
Ala	Val	Ala	Phe	Ile	Glu	Arg	Tyr	Phe	Gln	Ser	Tyr	Pro	Lys	Val	Arg	
	690					695					700					

Ala Trp Ile Glu Lys Thr Leu Ala Glu Gly Arg Glu Arg Gly Tyr Val
705 710 715 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Ala Ser Arg
725 730 735

Val Lys Ser Ile Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765

Phe Pro Arg Leu Gln Glu Leu Gly Ala Arg Met Leu Leu Gln Val His
770 775 780

Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Gln Ala Glu Glu Val Ala
785 790 795 800

Gln Glu Ala Lys Arg Thr Met Glu Glu Val Trp Pro Leu Lys Val Pro
805 810 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Ala
820 825 830

<210> 252
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 252
gccgccctcc tgaagcggct taaggg

26

<210> 253
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 253
cccttaagcc gcttcaggag ggcggc

26

<210> 254
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 254
atcggcaaga cgagaagac gggcaagc

28

<210> 255
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 255
gcttgcccgt cttctgcgtc ttgccgat 28

<210> 256
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 256
ttgcaggtgc acaacgaact ggtcctc 27

<210> 257
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 257
gaggaccagt tcgttgtgca cctgcaa 27

<210> 258
<211> 2526
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 258
atgaattcca cccactttt tgacctggag gaacccccca agcgggtgct tctggtggac 60
ggccaccacc tggcctaccg caccttctat gccctgagcc tcaccacctc ccgggggggag 120
ccggtgcaga tggctctacgg cttcgcccgg agcctcctca aggccttgaa ggaggacgga 180
caggcgggtgg tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
gcctacaagg cgggccgggc cccacccccg gaggacttcc cccgccagct cgccttggtc 300
aagcggctgg tggaccttct gggcctggtc cgcctcgagg ccccggggta cgaggcggac 360
gacgtcctgg gcaccctggc caagaaggcc gaaagggagg ggatggaggt gcgcatcctc 420
acgggagacc gggacttctt ccagctcctc tccgagaagg tctcggtcct cctgccggac 480

gggaccctgg tcaccccaaa ggacgtccag gagaagtacg gggtgccccc ggagcgctgg 540
 gtggacttcc gcgccctcac gggggaccgc tcggacaaca tccccggggg ggccggggata 600
 ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
 aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
 ctccacctct ccttagacct ggcccgcctc cgcaccgacc tccccctgga ggtggacttt 780
 aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840
 gagttcggaa gcctcctcca cgagttcggc ctctctggag gggagaagcc ccgggaggag 900
 gccccctggc ccccgccga aggggccttc gtgggcttcc tcctttcccg caaggagccc 960
 atgtggggcg agcttctggc cctggcgcg gcctcggcg gccgcgtgca ccgggcagca 1020
 gacccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc 1080
 gccgtcttgg cctcgaggga ggggctagac ctctgccccg gggacgaccc catgctcctc 1140
 gcctacctcc tggaccttc gaacaccacc cccgaggggg tggcgcgcg ctacgggggg 1200
 gagtggacgg aggacggcg ccaccggggc ctctctcgg agaggctcca tcggaacctc 1260
 cttaaagcgc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc 1320
 ctctcccggg tcctggccca tatggaggcc accggggtac ggcgggacgt ggctacctt 1380
 caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc 1440
 ttggcgggcc accccttcaa cctcaactcc cgggaccagc tggaaagggg gctctttgac 1500
 gagcttaggc ttccgcctt gaagaagacg aagaagacag gcaagcgctc caccagcgcc 1560
 gcggtgctgg aggccttac ggaggccac cccatcgctg agaagatcct ccagcaccgg 1620
 gagctacca agctcaagaa cacctacgtg gacccctcc caagcctcgt ccacccgagg 1680
 acgggcccgc tccacaccg cttcaaccag acggccacgg ccacggggag gcttagtagc 1740
 tccgaccca acctgcagaa catccccgtc cgcacccctc tgggcccagag gatccgccg 1800
 gccttcgtgg ccgaggcggt ttgggcgttg gtggcccttg actatagcca gatagagctc 1860
 cgcgtcctcg ccacctctc cggggacgaa aacctgatca gggcttcca ggaggggaag 1920
 gacatccaca ccagaccgc aagctggatg ttcggcgctc ccccgagggc cgtggacccc 1980
 ctgatgcgcc gggcgccaa gacggtgaac ttcggcgctc tctacggcat gtccgcccat 2040
 aggctctccc aggagcttgc catcccctac gaggagggcg tggcctttat agagcgctac 2100
 ttccaaagct tcccaaaggt gcgggccttg atagaaaaga ccctggagga ggggaggaag 2160
 cggggctacg tggaaacctt cttcggaaga aggcgctacg tgcccacact caacgcccgg 2220
 gtgaagagcg tcaggaggc cgcggagcg atggccttca acatgcccg ccagggcacc 2280
 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg 2340

gcccgcacatgc tcctccaggt cgccaacgag ctctctctgg agggccccca agcgcggggcc 2400
gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc 2460
ctggaggtgg aggtggggat gggggaggac tggctttccg ccaaggggtca ccaccaccac 2520
caccac 2526

<210> 259
<211> 842
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 259
Met Asn Ser Thr Pro Leu Phe Asp Leu Glu Glu Pro Pro Lys Arg Val
1 5 10 15
Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Tyr Ala Leu
20 25 30
Ser Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Met Val Tyr Gly Phe
35 40 45
Ala Arg Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Gln Ala Val Val
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80
Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg Leu
100 105 110
Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
115 120 125
Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
130 135 140
Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
145 150 155 160
Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
165 170 175
Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
180 185 190
Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
195 200 205
Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
210 215 220
Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
225 230 235 240

Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ala Ser Gly Gly Arg Val
 325 330 335
 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
 355 360 365
 Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
 405 410 415
 His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
 420 425 430
 Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
 435 440 445
 Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
 450 455 460
 Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
 465 470 475 480
 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
 485 490 495
 Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Lys Lys Thr Lys Lys
 500 505 510
 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
 515 520 525
 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
 530 535 540
 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
 545 550 555 560
 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
 565 570 575

Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
 580 585 590
 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
 595 600 605
 Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
 610 615 620
 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
 625 630 635 640
 Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
 645 650 655
 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
 660 665 670
 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
 675 680 685
 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
 690 695 700
 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
 705 710 715 720
 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
 725 730 735
 Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
 740 745 750
 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
 755 760 765
 Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
 770 775 780
 Leu Gln Val Ala Asn Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
 785 790 795 800
 Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
 805 810 815
 Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
 820 825 830
 Ser Ala Lys Gly His His His His His His
 835 840

<210> 260

<211> 2514

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 260

atgaattccc tgcccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60

ctggcctacc gtaccttttt tgcctgaag ggcctcacca ccagccgcgg ggagccggtc 120
caggcgggtgt acgggtttgc caagagcctt ttgaaggcgc taagggaaga cgggggatgtg 180
gtgatcgtgg tgtttgacgc caaggcccc tccttccgcc accagacctt cgaggcctac 240
aaggcggggc gggctccac ccccgaggac tttccccggc agcttgccct tatcaaggag 300
atggtggacc ttttgggcct ggagcgcctc gaggtgccgg gctttgaagc ggatgacgtc 360
ctggctaccc tggccaagaa ggcggaaaag gaaggctacg aagtgcgcat cctcaccgcg 420
gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggaggggttac 480
ctgatcacc cggagtggct ttgggagaag tatgggctta agccttccca gtgggtggac 540
taccgggcct tggccgggga cccttccgac aacatccccg gcgtgaaggg catcggggag 600
aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg 660
gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag 720
ctatccctgg agctatcccg ggtgcacacg gacttgctcc ttcaggtgga cttcgcccgg 780
cgccgggagc cggaccggga ggggcttaag gcctttttgg agaggctgga gttcggaagc 840
ctcctccacg agttcggcct gttggaaagc ccggtggcgg cggaggaagc tccctggccg 900
cccccgagg gagccttcgt ggggtacgtt ctttcccgcc ccgagcccat gtgggcggag 960
cttaacgcct tggccgccgc ctggggcggc cgctgcacc gggcagcaga ccccttggcg 1020
gggctaaagg acctcaagga ggtccggggc ctctcgcca aggacctcgc cgtcttggcc 1080
tcgagggagg ggctagacct cgtgcccggg gacgaccca tgctcctcgc ctacctctg 1140
gacccttcga acaccaccc cgaggggggtg gcgcggcgct acggggggga gtggacggag 1200
gacgcgccc accgggccct cctctcggag aggctccatc ggaacctcct taagcgcctc 1260
gagggggagg agaagctcct ttggctctac cacgaggtgg aaaagccctt ctcccgggtc 1320
ctggcccata tggaggccac cgggggtacg cgggacgtgg cctaccttca ggccctttcc 1380
ctggagcttg cggaggagat ccgccgcctc gaggaggagg tcttcgctt ggccggccac 1440
cccttcaacc tcaactcccg ggaccagctg gaaagggtgc tctttgacga gcttaggctt 1500
ccgccttgaa agaagacgaa gaagacaggc aagcgtcca ccagcgccgc ggtgctggag 1560
gccctacggg agggccaccc catcgtggag aagatcctcc agcaccggga gctaccaag 1620
ctcaagaaca cctacgtgga cccctccca agcctcgtcc acccgaggac gggccgcctc 1680
cacaccgct tcaaccagac ggccacggcc acggggaggc ttagtagctc cgacccaac 1740
ctgcagaaca tccccgtccg caccctcttg ggccagagga tccgcggggc cttcgtggcc 1800
gaggcgggtt gggcgttggg ggccctggac tatagccaga tagagctccg cgtcctcgcc 1860
cacctctccg gggacgaaa cctgatcagg gtcttcagg aggggaagga catccacacc 1920

cagaccgcaa gctggatgtt cggcgctccc cgggaggccg tggaccccct gatgcgccgg 1980
 ggggccaaga cgggtgaactt cggcgctcctc tacggcatgt ccgcccatag gctctcccag 2040
 gagcttgcca tcccctacga ggaggcggtg gcctttatag agcgctactt ccaaagcttc 2100
 cccaaggtgc gggcctggat agaaaagacc ctggaggagg ggaggaagcg gggctacgtg 2160
 gaaaccctct tcggaagaag gcgctacgtg cccgacctca acgcccgggt gaagagcgtc 2220
 agggaggccg cggagcgcat ggccttcaac atgcccgtcc agggcaccgc cgccgacctc 2280
 atgaagctcg ccatggtgaa gctcttcccc cgctccggg agatgggggc ccgcatgctc 2340
 ctccaggtcg ccaacgagct cctcctggag gcccccaag cgcgggccga ggaggtggcg 2400
 gctttggcca aggaggccat ggagaaggcc tatcccctcg ccgtgcccct ggaggtggag 2460
 gtggggatgg gggaggactg gctttccgcc aagggtcacc accaccacca ccac 2514

<210> 261

<211> 838

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 261

Met Asn Ser Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
 1 5 10 15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
 20 25 30

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
 35 40 45

Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Val Val Ile Val Val
 50 55 60

Phe Asp Ala Lys Ala Pro Ser Phe Arg His Gln Thr Tyr Glu Ala Tyr
 65 70 75 80

Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
 85 90 95

Leu Ile Lys Glu Met Val Asp Leu Leu Gly Leu Glu Arg Leu Glu Val
 100 105 110

Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala
 115 120 125

Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu
 130 135 140

Tyr Gln Leu Leu Ser Glu Arg Ile Ser Ile Leu His Pro Glu Gly Tyr
 145 150 155 160

Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys Pro Ser
 165 170 175

[illegible]

<210> 262
<211> 2526
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 262
atgaattcca cccactttt tgacctggag gaacccccca agcgggtgct tctggtggac 60
ggccaccacc tggcctaccg caccttctat gccctgagcc tcaccacctc ccggggggag 120
ccggtgcaga tggctctacg cttcgcccgg agcctcctca aggccttgaa ggaggacgga 180
caggcgggtgg tcgtggtctt tgacgccaag gccccctcct tccgccacga ggcctacgag 240
gcctacaagg cgggccgggc cccacccccg gaggacttcc cccgccagct cgccttggtc 300
aagcggctgg tggaccttct gggcctggtc cgcctcgagg ccccggggta cgaggcggac 360
gacgtcctgg gcaccctggc caagaaggcc gaaagggagg ggatggaggt gcgcacctc 420
acgggagacc gggacttctt ccagctcctc tccgagaagg tctcggtcct cctgccggac 480
gggaccctgg tcaccccaaa ggacgtccag gagaagtacg gggtgcccc ggagcgctgg 540
gtggacttcc gcgcctcac gggggaccgc tcggacaaca tccccggggt ggcggggata 600
ggggagaaga ccgcccttcg actcctcgca gagtggggga gcgtggaaaa cctcctgaag 660
aacctggacc gggtaaagcc ggactcgctc cggcgcaaga tagaggcgca cctcgaggac 720
ctccacctct ccttagacct ggcccgcatc cgcaccgacc tccccctgga ggtggacttt 780
aaggccctgc gccgcaggac ccccgacctg gagggcctga gggccttttt ggaggagctg 840
gagttcggaa gcctcctcca cgagttcggc ctctctggag gggagaagcc ccgggaggag 900
gccccctggc ccccgcccga aggggccttc gtgggcttcc tcctttcccg caaggagccc 960
atgtgggcgg agcttctggc cctggcggcg gcctcgggcg gccgcgtcca ccgggcccc 1020
gagccttata aagccctcag ggacctgaag gaggcgcggg ggcttctcgc caaagacctg 1080
agcgttctgg ccctgaggga aggccttggc ctcccgcccg gcgacgacct catgctcctc 1140
gcctacctcc tggacccttc gaacaccacc cccgaggggg tggcccgggc ctacggcggg 1200
gagtggacgg aggaggcggg ggagcgggccc gccctttccg agaggctctt cgccaacctg 1260
cttaagaggc ttgaggggga ggagaggctc ctttggtttt accgggaggt ggagaggccc 1320
ctttccgctg tcctggccca tatggaggcc acgggggtgc gcctggacgt ggcctatctc 1380
agggccttgt ccctggaggt ggccgaggag atcgcccgcc tcgaggccga ggtcttccgc 1440
ctggccggcc accccttcaa cctcaactcc cgggaccagc tggaaagggc cctctttgac 1500
gagctagggc ttcccgccat caagaagacg caaaagaccg gcaagcgctc caccagcgcc 1560

gccgtcctgg aggccctccg cgaggccac cccatcgtgg agaagatcct gcagtaccgg 1620
gagctcacca agctgaagag cacctacatt gaccccttgc cggacctcat ccaccccagg 1680
acggggccgcc tccacacccg cttcaaccag acggccacgg ccacggggcag gctaagtagc 1740
tccgatccca acctccagaa catccccgtc cgcaccccg cgtgggcagag gatccgccgg 1800
gccttcatcg ccgaggaggg gtggctattg gtggccctgg actatagcca gatagagctc 1860
aggggtgctgg cccacctctc cggcgacgag aacctgatcc ggggtcttcca ggagggggcgg 1920
gacatccaca cggagaccgc cagctggatg ttcggcgctcc cccgggaggc cgtggacccc 1980
ctgatgcgcc gggcgggcaa gaccatcaac ttcgggggtcc tctacggcat gtcggccccc 2040
cgcctctccc aggagctagc catcccttac gaggaggccc aggccttcat tgagcgctac 2100
tttcagagct tccccaaggt gcgggcctgg attgagaaga ccctggagga gggcaggagg 2160
cgggggtacg tggagaccct cttcgccgc cgcgctacg tgccagacct agaggcccgg 2220
gtgaagagcg tgcgggaggc ggccgagcgc atggccttca acatgcccgt ccagggcacc 2280
gccgccgacc tcatgaagct ggctatggtg aagctcttcc ccaggctgga ggaaatgggg 2340
gccaggatgc tccttcaggt cgccaacgag ctggtcctcg agggcccaaa agagagggcg 2400
gaggccgtgg cccggctggc caaggaggtc atggaggggg tgtatccctt ggccgtgccc 2460
ctggaggtgg aggtggggat aggggaggac tggctctccg ccaaggagca ccaccaccac 2520
caccac 2526

<210> 263

<211> 842

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 263

Met	Asn	Ser	Thr	Pro	Leu	Phe	Asp	Leu	Glu	Glu	Pro	Pro	Lys	Arg	Val
1				5				10					15		
Leu	Leu	Val	Asp	Gly	His	His	Leu	Ala	Tyr	Arg	Thr	Phe	Tyr	Ala	Leu
			20					25					30		
Ser	Leu	Thr	Thr	Ser	Arg	Gly	Glu	Pro	Val	Gln	Met	Val	Tyr	Gly	Phe
		35				40					45				
Ala	Arg	Ser	Leu	Leu	Lys	Ala	Leu	Lys	Glu	Asp	Gly	Gln	Ala	Val	Val
	50				55					60					
Val	Val	Phe	Asp	Ala	Lys	Ala	Pro	Ser	Phe	Arg	His	Glu	Ala	Tyr	Glu
65				70					75					80	
Ala	Tyr	Lys	Ala	Gly	Arg	Ala	Pro	Thr	Pro	Glu	Asp	Phe	Pro	Arg	Gln
			85					90						95	

Leu Ala Leu Val Lys Arg Leu Val Asp Leu Leu Gly Leu Val Arg Leu
 100 105 110
 Glu Ala Pro Gly Tyr Glu Ala Asp Asp Val Leu Gly Thr Leu Ala Lys
 115 120 125
 Lys Ala Glu Arg Glu Gly Met Glu Val Arg Ile Leu Thr Gly Asp Arg
 130 135 140
 Asp Phe Phe Gln Leu Leu Ser Glu Lys Val Ser Val Leu Leu Pro Asp
 145 150 155 160
 Gly Thr Leu Val Thr Pro Lys Asp Val Gln Glu Lys Tyr Gly Val Pro
 165 170 175
 Pro Glu Arg Trp Val Asp Phe Arg Ala Leu Thr Gly Asp Arg Ser Asp
 180 185 190
 Asn Ile Pro Gly Val Ala Gly Ile Gly Glu Lys Thr Ala Leu Arg Leu
 195 200 205
 Leu Ala Glu Trp Gly Ser Val Glu Asn Leu Leu Lys Asn Leu Asp Arg
 210 215 220
 Val Lys Pro Asp Ser Leu Arg Arg Lys Ile Glu Ala His Leu Glu Asp
 225 230 235 240
 Leu His Leu Ser Leu Asp Leu Ala Arg Ile Arg Thr Asp Leu Pro Leu
 245 250 255
 Glu Val Asp Phe Lys Ala Leu Arg Arg Arg Thr Pro Asp Leu Glu Gly
 260 265 270
 Leu Arg Ala Phe Leu Glu Glu Leu Glu Phe Gly Ser Leu Leu His Glu
 275 280 285
 Phe Gly Leu Leu Gly Gly Glu Lys Pro Arg Glu Glu Ala Pro Trp Pro
 290 295 300
 Pro Pro Glu Gly Ala Phe Val Gly Phe Leu Leu Ser Arg Lys Glu Pro
 305 310 315 320
 Met Trp Ala Glu Leu Leu Ala Leu Ala Ala Ala Ser Gly Gly Arg Val
 325 330 335
 His Arg Ala Pro Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala
 340 345 350
 Arg Gly Leu Leu Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly
 355 360 365
 Leu Gly Leu Pro Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
 370 375 380
 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
 385 390 395 400
 Glu Trp Thr Glu Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu
 405 410 415
 Phe Ala Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp
 420 425 430

Met	Val	Lys	Leu	Phe	Pro	Arg	Leu	Glu	Glu	Met	Gly	Ala	Arg	Met	Leu
770						775					780				
Leu	Gln	Val	Ala	Asn	Glu	Leu	Val	Leu	Glu	Ala	Pro	Lys	Glu	Arg	Ala
785					790					795					800
Glu	Ala	Val	Ala	Arg	Leu	Ala	Lys	Glu	Val	Met	Glu	Gly	Val	Tyr	Pro
				805					810					815	
Leu	Ala	Val	Pro	Leu	Glu	Val	Glu	Val	Gly	Ile	Gly	Glu	Asp	Trp	Leu
			820					825					830		
Ser	Ala	Lys	Glu	His	His	His	His	His	His						
		835					840								

<210> 264
 <211> 2514
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic

<400> 264
 atgaattccc tgcccctctt tgagcccaag ggccgggtgc ttctggtgga cggccaccac 60
 ctggcctacc gtaccttttt tgcctgaag ggcctcacca ccagccgagg ggagccggtc 120
 caggcgggtgt acgggttttg caagagcctt ttgaaggcgc taagggaaga cggggatgtg 180
 gtgatcgtgg tgtttgacgc caaggcccc tccttcgcc accagacctt cgaggcctac 240
 aaggcggggc gggctccac ccccgaggac tttcccggc agcttgccct tatcaaggag 300
 atggtggacc ttttgggcct ggagcgcctc gaggtgccgg gctttgaagc ggatgacgtc 360
 ctggctaccc tggccaagaa ggcgaaaag gaaggctacg aagtgcgcat cctcaccgcg 420
 gaccgggacc tttaccagct tctttcggag cgaatctcca tccttcaccc ggaggggttac 480
 ctgatcacc cggagtggct ttgggagaag tatgggctta agccttccca gtgggtggac 540
 taccgggcct tggccgggga cccttcgac aacatccccg gcgtgaaggg catcggggag 600
 aagacggcgg ccaagctgat ccgggagtgg ggaagcctgg aaaaccttct taagcacctg 660
 gaacaggtga aacctgcctc cgtgcgggag aagatcctta gccacatgga ggacctcaag 720
 ctatccctgg agctatccc ggtgcacacg gacttgctcc ttcaggtgga cttcgcccgg 780
 cgccgggagc cggaccggga ggggcttaag gccttttttg agaggctgga gttcggaagc 840
 ctctccacg agttcggcct gttggaaagc ccggtggcgg cggaggaagc tccctggccg 900
 cccccgagg gagccttcgt ggggtacgtt ctttcccgc ccgagcccat gtgggaggag 960
 cttaacgcct tggccgccgc ctggggcggc cgcgtccacc gggccccga gccttataaa 1020
 gccctcaggg acctgaagga ggcgcggggg cttctcgcca aagacctgag cgttctggcc 1080
 ctgagggaag gccttggcct ccgcccggc gacgaccca tgctctcgc ctacctcctg 1140

gacccttcga acaccacccc cgagggggtg gcccggcgct acggcgggga gtggacggag 1200
gaggcggggg agcgggccgc cttttccgag aggtcttcg ccaacctgct taagaggctt 1260
gagggggagg agaggctcct ttggctttac cgggaggtgg agaggcccct ttccgctgtc 1320
ctggcccata tggaggccac gggggtgcgc ctggacgtgg cctatctcag ggccttgtcc 1380
ctggaggtgg ccgaggagat cggccgcctc gagggcgagg tcttcgcctt ggccggccac 1440
cccttcaacc tcaactcccg ggaccagctg gaaagggtcc tctttgacga gctagggtt 1500
cccgccatca agaagacgca aaagaccggc aagcgctcca ccagcgccgc cgtcctggag 1560
gccctccgag agggccaccc catcgtggag aagatcctgc agtaccggga gctcaccaag 1620
ctgaagagca cctacattga ccccttgccg gacctcatcc accccaggac gggccgcctc 1680
cacaccgct tcaaccagac ggccacggcc acgggcaggc taagtagctc cgatcccaac 1740
ctccagaaca tccccgtccg caccgcctt gggcagagga tccgcccggc cttcatcgcc 1800
gaggaggggt ggctattggt ggccctggac tatagccaga tagagctcag ggtgctggcc 1860
cacctctccg gcgacgagaa cctgatccgg gtcttccagg agggcgggga catccacacg 1920
gagaccgcca gctggatggt cggcgctccc cgggagggcg tggaccccct gatgcgccgg 1980
gcgggccaaga ccatcaactt cggggtcctc tacggcatgt cggcccaccg cctctcccag 2040
gagctagcca tcccttacga ggaggcccag gccttcattg agcgctactt tcagagcttc 2100
cccaaggtgc gggcctggat tgagaagacc ctggaggagg gcaggaggcg ggggtacgtg 2160
gagaccctct tcggccgccg ccgctacgtg ccagacctag agggccgggt gaagagcgtg 2220
cgggagggcg ccgagcgcat ggccttcaac atgcccgtcc agggcaccgc cgccgacctc 2280
atgaagctgg ctatggtgaa gctcttcccc aggtctggagg aaatgggggc caggatgctc 2340
cttcaggtcg ccaacgagct ggtcctcgag gcccctaaag agagggcgga ggccgtggcc 2400
cggtggcca aggaggtcat ggagggggtg tatcccctgg ccgtgccctt ggaggtggag 2460
gtggggatag gggaggactg gctctccgcc aaggagcacc accaccacca ccac 2514

<210> 265

<211> 838

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 265

Met Asn Ser Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1 5 10 15

Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
20 25 30

Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
 35 40 45
 Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Val Val Ile Val Val
 50 55 60
 Phe Asp Ala Lys Ala Pro Ser Phe Arg His Gln Thr Tyr Glu Ala Tyr
 65 70 75 80
 Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
 85 90 95
 Leu Ile Lys Glu Met Val Asp Leu Leu Gly Leu Glu Arg Leu Glu Val
 100 105 110
 Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala
 115 120 125
 Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu
 130 135 140
 Tyr Gln Leu Leu Ser Glu Arg Ile Ser Ile Leu His Pro Glu Gly Tyr
 145 150 155 160
 Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys Pro Ser
 165 170 175
 Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile
 180 185 190
 Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Ala Lys Leu Ile Arg
 195 200 205
 Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys His Leu Glu Gln Val Lys
 210 215 220
 Pro Ala Ser Val Arg Glu Lys Ile Leu Ser His Met Glu Asp Leu Lys
 225 230 235 240
 Leu Ser Leu Glu Leu Ser Arg Val His Thr Asp Leu Leu Leu Gln Val
 245 250 255
 Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Lys Ala Phe
 260 265 270
 Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
 275 280 285
 Glu Ser Pro Val Ala Ala Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
 290 295 300
 Ala Phe Val Gly Tyr Val Leu Ser Arg Pro Glu Pro Met Trp Ala Glu
 305 310 315 320
 Leu Asn Ala Leu Ala Ala Ala Trp Gly Gly Arg Val His Arg Ala Pro
 325 330 335
 Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu
 340 345 350
 Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro
 355 360 365

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val
705 710 715 720

Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg
725 730 735

Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765

Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val Ala
770 775 780

Asn Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala
785 790 795 800

Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
805 810 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

His His His His His His
835

<210> 266
<211> 2505
<212> DNA
<213> Thermus thermophilus

<400> 266
atggaggcga tgcttccgct ctttgaaccc aaaggccggg tcctcctggt ggacggccac 60
cacctggcct accgcacctt ctctgccctg aagggcctca ccacgagccg gggcgaaccg 120
gtgcaggcgg tctacggctt cgccaagagc ctctcaagg ccctgaagga ggacgggtac 180
aaggccgtct tcgtggtctt tgacgccaaag gccccctcct tccgccacga ggcctacgag 240
gcctacaagg cggggagggc cccgaccccc gaggacttcc cccggcagct cgcctcatc 300
aaggagctgg tggacctcct ggggtttacc cgcctcgagg tccccggcta cgaggcggac 360
gacgttctcg ccaccctggc caagaaggcg gaaaaggagg ggtacgaggt gcgcacctc 420
accgccgacc gcgacctcta ccaactcgtc tccgaccgag tcgccgtcct ccaccccgag 480
ggccaccta tcaccccgga gtggcttttg gagaagtacg gcctcaggcc ggagcagtgg 540
gtggacttcc gcgcctcgt gggggacccc tccgacaacc tccccggggt caagggcatc 600
ggggagaaga ccgccctcaa gtcctcaag gagggggaa gcctggaaaa cctcctcaag 660
aacctggacc gggtaaagcc agaaaacgtc cgggagaaga tcaaggccca cctggaagac 720
ctcaggctct ccttgagct ctcccgggtg cgcaccgacc tccccctgga ggtggacctc 780
gcccaggggc gggagcccga ccgggagggg cttagggcct tcctggagag gctggagttc 840

ggagcctcc tccacgagtt cggcctcctg gagggccccg cccccctgga ggaggcccc 900
 tggccccgc cggaaggggc cttcgtgggc ttcgtcctct cccgccccga gcccatgtgg 960
 gcggagctta aagccctggc cgctgcagg gacggccggg tgcaccgggc agcagacccc 1020
 ttggcggggc taaaggacct caaggaggtc cggggcctcc tcgccaagga cctcgccgtc 1080
 ttggcctcga gggaggggct agacctcgtg cccggggacg accccatgct cctcgccctac 1140
 ctcttgacc cctccaacac ccccccgag ggggtggcgc ggcgctacgg gggggagtgg 1200
 acggaggacg ccgcccaccg ggccctcctc tcggagaggc tccatcgaa cctccttaag 1260
 cgctcgagg gggaggagaa gtccttttgg ctctaccag aggtggaaaa gccctctctc 1320
 cgggtcctgg ccacatgga ggccaccggg gtacggctgg acgtggccta cttcaggcc 1380
 ctttccctgg agcttgcgga ggagatccgc cgcctcgagg aggaggtctt ccgcttggcg 1440
 ggccaccct tcaacctcaa ctcccgggac cagctggaaa ggggtgctctt tgacgagctt 1500
 aggcttccc ccttggggaa gacgcaaaag acaggcaagc gctccaccag cgcgcggtg 1560
 ctggaggccc tacgggaggc ccaccccatc gtggagaaga tcctccagca ccgggagctc 1620
 accaagctca agaacaccta cgtggacccc ctccaagcc tcgtccacc gaggacgggc 1680
 cgctccaca ccgcttcaa ccagacggcc acggccaccg ggaggcttag tagctccgac 1740
 cccaacctgc agaacatccc cgtccgcacc cccttgggccc agaggatccg ccgggccttc 1800
 gtggccgagg cgggttgggc gttggtggcc ctggactata gccagataga gctccgcgtc 1860
 ctgcccacc tctccgggga cgaaaacctg atcagggtct tccaggagg gaaggacatc 1920
 cacaccaga ccgcaagctg gatgttcggc gtccccccg aggcctgga cccctgatg 1980
 cgccgggagg ccaagacggt gaacttcggc gtctctacg gcatgtccgc ccataggctc 2040
 tcccaggagc ttgccatccc ctacgaggag gcggtggcct ttatagagcg ctacttccaa 2100
 agcttcccca aggtgcgggc ctggatagaa aagaccctgg aggaggggag gaagcggggc 2160
 tacgtggaac ccctcttcgg aagaaggcgc tacgtgcccg acctcaacgc ccgggtgaag 2220
 agcgtcaggg aggcgcgga gcgcatggcc ttcaacatgc ccgtccaggg caccgcgcgc 2280
 gacctcatga agctcgccat ggtgaagctc ttcccccgcc tccgggagat gggggccgc 2340
 atgctcctcc aggtccacga cgagtcctc ctggaggccc cccaagcgcg ggccgaggag 2400
 gtggcggtt tggccaagga ggccatggag aaggcctatc ccctcgccgt gccctggag 2460
 gtggaggtg ggatggggga ggactggctt tccgccaagg gttag 2505

<210> 267

<211> 834

<212> PRT

<213> *Thermus aquaticus*

<400> 267
Met Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
1 5 10 15
Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly
20 25 30
Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
35 40 45
Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80
Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
115 120 125
Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
130 135 140
Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145 150 155 160
Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
195 200 205
Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg
210 215 220
Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp
225 230 235 240
Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu
245 250 255
Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly Leu Arg
260 265 270
Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
275 280 285
Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro Pro Pro
290 295 300
Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp
305 310 315 320
Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val His Arg

										325				330				335			
Ala	Ala	Asp	Pro	Leu	Ala	Gly	Leu	Lys	Asp	Leu	Lys	Glu	Val	Arg	Gly						
			340				345						350								
Leu	Leu	Ala	Lys	Asp	Leu	Ala	Val	Leu	Ala	Ser	Arg	Glu	Gly	Leu	Asp						
		355				360						365									
Leu	Val	Pro	Gly	Asp	Asp	Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro						
		370				375						380									
Ser	Asn	Thr	Thr	Pro	Glu	Gly	Val	Ala	Arg	Arg	Tyr	Gly	Gly	Glu	Trp						
				390				395						400							
Thr	Glu	Asp	Ala	Ala	His	Arg	Ala	Leu	Leu	Ser	Glu	Arg	Leu	His	Arg						
				405				410						415							
Asn	Leu	Leu	Lys	Arg	Leu	Glu	Gly	Glu	Glu	Lys	Leu	Leu	Trp	Leu	Tyr						
			420				425						430								
His	Glu	Val	Glu	Lys	Pro	Leu	Ser	Arg	Val	Leu	Ala	His	Met	Glu	Ala						
		435				440						445									
Thr	Gly	Val	Arg	Leu	Asp	Val	Ala	Tyr	Leu	Gln	Ala	Leu	Ser	Leu	Glu						
		450				455						460									
Leu	Ala	Glu	Glu	Ile	Arg	Arg	Leu	Glu	Glu	Glu	Val	Phe	Arg	Leu	Ala						
				470				475						480							
Gly	His	Pro	Phe	Asn	Leu	Asn	Ser	Arg	Asp	Gln	Leu	Glu	Arg	Val	Leu						
				485				490						495							
Phe	Asp	Glu	Leu	Arg	Leu	Pro	Ala	Leu	Gly	Lys	Thr	Gln	Lys	Thr	Gly						
			500				505						510								
Lys	Arg	Ser	Thr	Ser	Ala	Ala	Val	Leu	Glu	Ala	Leu	Arg	Glu	Ala	His						
		515				520						525									
Pro	Ile	Val	Glu	Lys	Ile	Leu	Gln	His	Arg	Glu	Leu	Thr	Lys	Leu	Lys						
		530				535						540									
Asn	Thr	Tyr	Val	Asp	Pro	Leu	Pro	Ser	Leu	Val	His	Pro	Arg	Thr	Gly						
				545				550						555							
Arg	Leu	His	Thr	Arg	Phe	Asn	Gln	Thr	Ala	Thr	Ala	Thr	Gly	Arg	Leu						
				565				570						575							
Ser	Ser	Ser	Asp	Pro	Asn	Leu	Gln	Asn	Ile	Pro	Val	Arg	Thr	Pro	Leu						
			580				585						590								
Gly	Gln	Arg	Ile	Arg	Arg	Ala	Phe	Val	Ala	Glu	Ala	Gly	Trp	Ala	Leu						
		595				600						605									
Val	Ala	Leu	Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Val	Leu	Ala	His	Leu						
		610				615						620									
Ser	Gly	Asp	Glu	Asn	Leu	Ile	Arg	Val	Phe	Gln	Glu	Gly	Lys	Asp	Ile						
		625				630						635			640						
His	Thr	Gln	Thr	Ala	Ser	Trp	Met	Phe	Gly	Val	Pro	Pro	Glu	Ala	Val						
				645				650						655							

Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
 660 665 670
 Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
 675 680 685
 Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
 690 695 700
 Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
 705 710 715 720
 Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
 725 730 735
 Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
 740 745 750
 Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
 755 760 765
 Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
 770 775 780
 Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
 785 790 795 800
 Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
 805 810 815
 Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
 820 825 830
 Lys Gly

<210> 268

<211> 832

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 268

Met Glu Phe Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
 1 5 10 15
 Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
 20 25 30
 Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
 35 40 45
 Ser Leu Leu Lys Ala Leu Arg Glu Asp Gly Asp Val Val Ile Val Val
 50 55 60
 Phe Asp Ala Lys Ala Pro Ser Phe Arg His Gln Thr Tyr Glu Ala Tyr
 65 70 75 80
 Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala

Leu Glu Arg Leu Lys Gly Glu Glu Arg Leu Leu Trp Leu Tyr Glu Glu
 420 425 430
 Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala Thr Gly
 435 440 445
 Val Arg Leu Asp Val Ala Tyr Leu Lys Ala Leu Ser Leu Glu Val Glu
 450 455 460
 Ala Glu Ile Arg Arg Phe Glu Glu Glu Val His Arg Leu Ala Gly His
 465 470 475 480
 Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Ile Phe Asp
 485 490 495
 Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg
 500 505 510
 Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile
 515 520 525
 Val Asp Arg Ile Leu Gln Tyr Arg Glu Leu Ser Lys Leu Lys Gly Thr
 530 535 540
 Tyr Ile Asp Pro Leu Pro Ala Leu Val His Pro Lys Thr Asn Arg Leu
 545 550 555 560
 His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser
 565 570 575
 Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln
 580 585 590
 Arg Ile Arg Arg Ala Phe Val Ala Glu Glu Gly Trp Arg Leu Val Val
 595 600 605
 Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly
 610 615 620
 Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Gln Asp Ile His Thr
 625 630 635 640
 Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val Asp Ser
 645 650 655
 Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly
 660 665 670
 Met Ser Ala His Arg Leu Ser Gly Glu Leu Ala Ile Pro Tyr Glu Glu
 675 680 685
 Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Tyr Pro Lys Val Arg
 690 695 700
 Ala Trp Ile Glu Lys Thr Leu Ala Glu Gly Arg Glu Arg Gly Tyr Val
 705 710 715 720
 Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Ala Ser Arg
 725 730 735
 Val Lys Ser Ile Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
 740 745 750

Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765

Phe Pro Arg Leu Gln Glu Leu Gly Ala Arg Met Leu Leu Gln Val His
770 775 780

Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Gln Ala Glu Glu Val Ala
785 790 795 800

Gln Glu Ala Lys Arg Thr Met Glu Glu Val Trp Pro Leu Lys Val Pro
805 810 815

Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Ala
820 825 830

<210> 269
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 269
gcgcggcggc ggggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccttag 60
cgc 63

<210> 270
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 270
gcgctagggc gctggcaagt gtagcgggtca 30

<210> 271
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 271
gatcgctgcg cgtaaccacc acacccgcgc cgcgc 35

<210> 272
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 272
ggactctgcc tcaagacggt agtcaacgtg 30

<210> 273
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 273
cacgttgact accgtc 16

<210> 274
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 274
catgtcaagc agtcctaact ttgaggcaga gtcc 34

<210> 275
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 275
cagaccatga attccacccc actttttgac ctggag 36

<210> 276
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 276
gtggacgcgg ccgcccgcgg ccgcccgcgg ggccag 36

<210> 277
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 277
cagaccatga attccctgcc cctctttgag cccaag

36

<210> 278
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 278
gtaaaccgcg ccgccccagg cggcggccaa ggcgtt

36

CLAIMS

We claim:

5 1. A composition comprising an enzyme, wherein said enzyme comprises a heterologous functional domain, wherein said heterologous functional domain provides altered functionality in a nucleic acid cleavage assay.

10 2. The composition of Claim 1, wherein said enzyme comprises a 5' nuclease.

 3. The composition of Claim 2, wherein said 5' nuclease comprises a thermostable 5' nuclease.

 4. The composition of Claim 1, wherein said enzyme comprises a polymerase.

 5. The composition of Claim 4, wherein said polymerase is altered in sequence relative to a naturally occurring sequence of a polymerase such that it exhibits reduced DNA synthetic activity from that of the naturally occurring polymerase.

 6. The composition of Claim 4, wherein said polymerase comprises a thermostable polymerase.

 7. The composition of Claim 6, wherein said thermostable polymerase comprises a polymerase from a *Thermus* species.

 8. The composition of Claim 7, wherein said *Thermus* species is selected from *Thermus aquaticus*, *Thermus flavus*, *Thermus thermophilus*, *Thermus filiformus*, and *Thermus scotoductus*.

30 9. The composition of Claim 1, wherein said heterologous functional domain

comprises an amino acid sequence that provides an improved nuclease activity in said nucleic acid cleavage assay.

10. The composition of Claim 1, wherein said heterologous functional domain
comprises an amino acid sequence that provides an improved substrate binding activity in said
nucleic acid cleavage assay.

11. The composition of Claim 1, wherein said heterologous functional domain
comprises an amino acid sequence that provides improved background specificity in said
nucleic acid cleavage assay.

12. The composition of Claim 1, wherein said heterologous functional domain
comprises two or more amino acids from a polymerase domain of a polymerase.

13. The composition of Claim 12, wherein at least one of said two or more amino
acids is from a palm region of said polymerase domain.

14. The composition of Claim 12, wherein at least one of said two or more amino
acids is from a thumb region of said polymerase domain.

15. The composition of Claim 12, wherein said polymerase comprises *Thermus
thermophilus* polymerase.

16. The composition of Claim 12, wherein said two or amino acids from said
polymerase domain comprise two or more amino acids from amino acids 300-650 of SEQ ID
NO:267.

17. The composition of Claim 1, wherein said enzyme comprises an amino acid
sequence selected from the group consisting of SEQ ID NOs:75, 77, 79, 81, 83, 85, 87, 89,
91, 93, 95, 97, 99, 101, 103, 106, 109, 112, 115, 118, 121, 124, 127, 130, 133, 136, 139,

142, 145, 148, 150, 153, 157, 160, 163, 166, 169, 172, 175, 178, 181, 184, 187, 190, 200, 202, 204, 206, 212, 214, 216, 218, 221, 226, 228, 230, 232, 234, 236, 239, 241, 243, 251, 259, 261, 263, and 265.

5 18. The composition of Claim 1, wherein said nucleic acid cleavage assay comprises cleavage of a DNA member of a substrate containing at least one RNA component.

 19. The composition of Claim 1, wherein said nucleic acid cleavage assay comprises an invasive cleavage assay.

10 20. A composition comprising a nucleic acid encoding the enzyme of Claim 1.

 21. The composition of Claim 20, wherein said nucleic acid is selected from the group consisting of SEQ ID NOs:74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 105, 108, 111, 114, 117, 120, 123, 126, 129, 132, 135, 138, 141, 144, 147, 149, 152, 156, 159, 162, 165, 168, 171, 174, 177, 180, 183, 186, 189, 199, 201, 203, 205, 211, 213, 215, 217, 220, 225, 227, 229, 231, 233, 235, 238, 240, 242, 250, 258, 260, 262, and 264.

 22. The composition of Claim 20, further comprising an expression vector operably linked to said nucleic acid.

 23. A composition comprising a host cell containing the composition of Claim 22.

25 24. A method for producing an altered enzyme with improved functionality in a nucleic acid cleavage assay comprising:

- a) providing an enzyme and a nucleic acid test substrate;
- b) introducing a heterologous functional domain into said enzyme to produce an altered enzyme;
- c) contacting said altered enzyme with said nucleic acid test substrate to produce cleavage products; and

d) detecting said cleavage products.

25. The method of Claim 24, wherein said introducing a heterologous functional domain comprises mutating one or more amino acids of said enzyme.

5

26. The method of Claim 24, wherein said introducing a heterologous functional domain into said enzyme comprises adding a functional domain from a protein into said enzyme.

10

27. The method Claim 26, wherein said adding a functional domain from a protein into said enzyme comprising removing a portion of said enzyme sequence prior to adding said functional domain of said protein.

15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160
165
170
175
180
185
190
195
200
205
210
215
220
225
230
235
240
245
250
255
260
265
270
275
280
285
290
295
300
305
310
315
320
325
330
335
340
345
350
355
360
365
370
375
380
385
390
395
400
405
410
415
420
425
430
435
440
445
450
455
460
465
470
475
480
485
490
495
500

28. The method of Claim 24, wherein said nucleic acid test substrate comprises a cleavage structure.

29. The method of Claim 28, wherein said cleavage structure comprises an RNA target nucleic acid.

30. The method of Claim 28, wherein said cleavage structure comprises an invasive cleavage structure.

31. The method of Claim 24, wherein said enzyme comprises a 5' nuclease.

32. The method of Claim 31, wherein said 5' nuclease comprises a thermostable 5' nuclease.

33. The method of Claim 24, wherein said enzyme comprises a polymerase.

34. The method of Claim 33, wherein said polymerase is altered in sequence

30

relative to a naturally occurring sequence of a polymerase such that it exhibits reduced DNA synthetic activity from that of the naturally occurring polymerase.

35. The method of Claim 34, wherein said polymerase comprises a thermostable polymerase.

36. The method of Claim 35, wherein said thermostable polymerase comprises a polymerase from a *Thermus* species.

37. The method of Claim 36, wherein said *Thermus* species is selected from *Thermus aquaticus*, *Thermus flavus*, *Thermus thermophilus*, *Thermus filiformus*, and *Thermus scotoductus*.

38. The method of Claim 24, wherein said heterologous functional domain comprises an amino acid sequence that provides an improved nuclease activity in said nucleic acid cleavage assay.

39. The method of Claim 24, wherein said heterologous functional domain comprises an amino acid sequence that provides an improved substrate binding activity in said nucleic acid cleavage assay.

40. The method of Claim 24, wherein said heterologous functional domain comprises an amino acid sequence that provides improved background specificity in said nucleic acid cleavage assay.

41. The method of Claim 24, wherein said heterologous functional domain comprises two or more amino acids from a polymerase domain of a polymerase.

42. The method of Claim 41, wherein at least one of said two or more amino acids is from a palm region of said polymerase domain.

43. The method of Claim 41, wherein at least one of said two or more amino acids is from a thumb region of said polymerase domain.

44. The method of Claim 41, wherein said polymerase comprises *Thermus thermophilus* polymerase.

45. The method of Claim 41, wherein said two or amino acids from said polymerase domain comprise two or more amino acids from amino acids 300-650 of SEQ ID NO:267.

46. The method of Claim 24, wherein said enzyme comprises an amino acid sequence selected from the group consisting of SEQ ID NOs:75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 106, 109, 112, 115, 118, 121, 124, 127, 130, 133, 136, 139, 142, 145, 148, 150, 153, 157, 160, 163, 166, 169, 172, 175, 178, 181, 184, 187, 190, 200, 202, 204, 206, 212, 214, 216, 218, 221, 226, 228, 230, 232, 234, 236, 239, 241, 243, 251, 259, 261, 263, and 265.

47. The method of Claim 24, wherein said altered enzyme comprises an amino acid sequence selected from the group consisting of SEQ ID NOs:75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 106, 109, 112, 115, 118, 121, 124, 127, 130, 133, 136, 139, 142, 145, 148, 150, 153, 157, 160, 163, 166, 169, 172, 175, 178, 181, 184, 187, 190, 200, 202, 204, 206, 212, 214, 216, 218, 221, 226, 228, 230, 232, 234, 236, 239, 241, 243, 251, 259, 261, 263, and 265.

48. An altered enzyme produced by the method of Claim 24.

49. A kit comprising the altered enzyme of Claim 48.

50. A kit comprising the composition of Claim 1.

51. The kit of Claim 50, further comprising at least one nucleic acid cleavage substrate.

52. The kit of Claim 51, further comprising at least one RNA capable of hybridizing to said at nucleic acid cleavage substrate.

53. The kit of Claim 50, further comprising a labeled oligonucleotide.

54. The kit of Claim 50, further comprising an invasive oligonucleotide.

55. A method for cleaving a nucleic acid comprising:

a) providing:

i) the enzyme of Claim 1; and

ii) a substrate nucleic acid; and

b) exposing said substrate nucleic acid to said enzyme.

56. The method of Claim 55, wherein said exposing said substrate nucleic acid to said enzyme produces at least one cleavage product.

57. The method of Claim 56, further comprising the step of c) detecting said cleavage product.

ABSTRACT

The present invention relates to novel enzymes designed for direct detection, characterization and quantitation of nucleic acids, particularly RNA. The present invention provides enzymes that recognize specific nucleic acid cleavage structures formed on a target RNA sequence and that cleave the nucleic acid cleavage structure in a site-specific manner to produce non-target cleavage products. The present invention provides enzymes having an improved ability to specifically cleave a DNA member of a complex comprising DNA and RNA nucleic acid strands.

10

Continued on next page

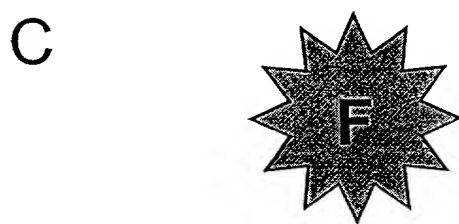
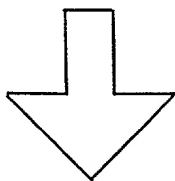
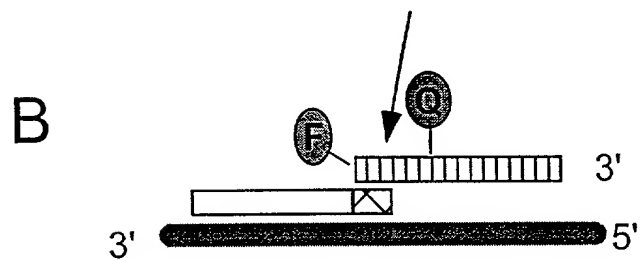
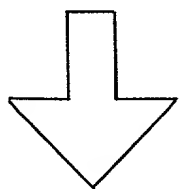
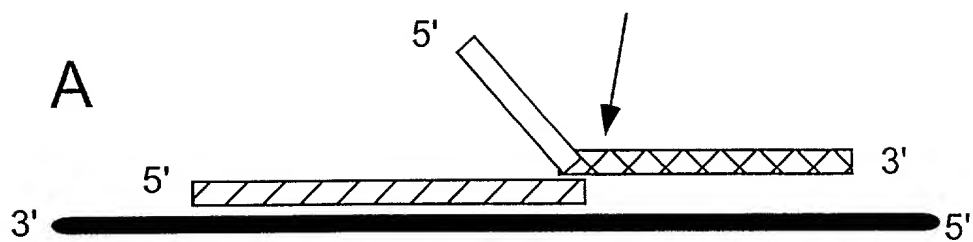
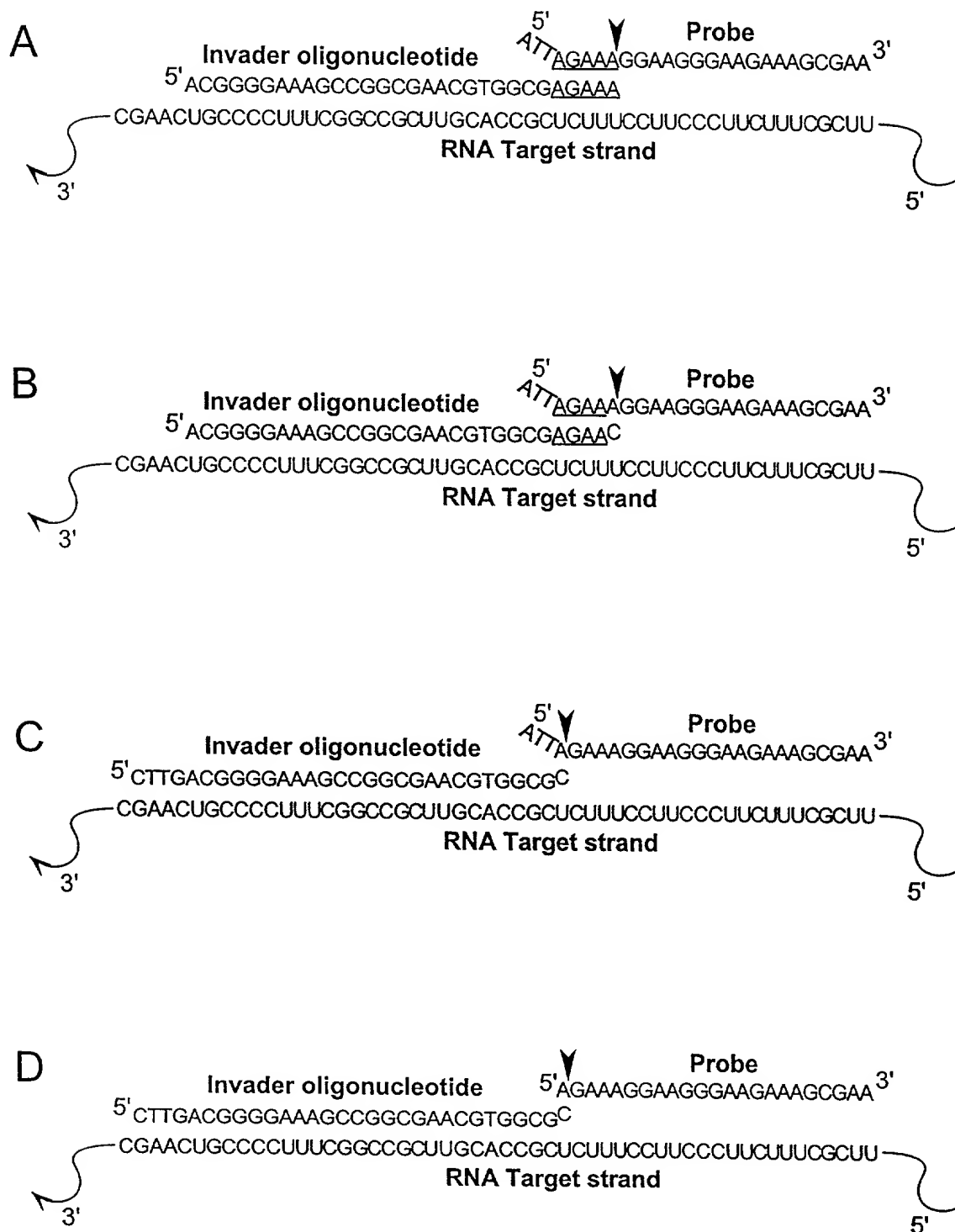


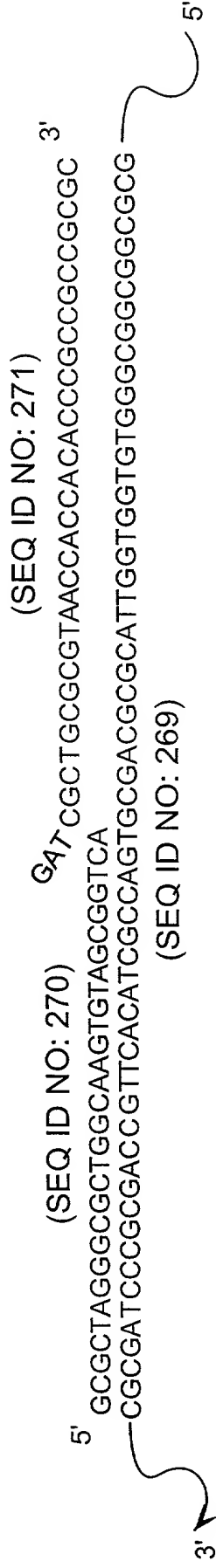
FIGURE 1

004260" 40622560

FIGURE 2



A



B

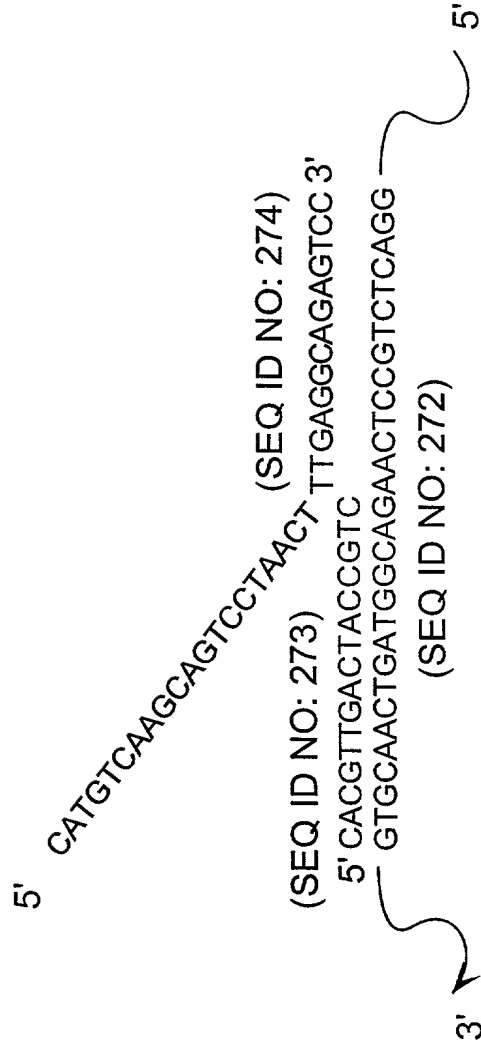


FIGURE 3

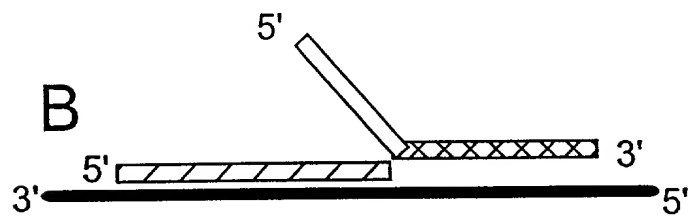
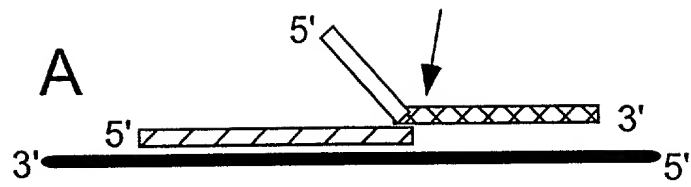


FIGURE 4

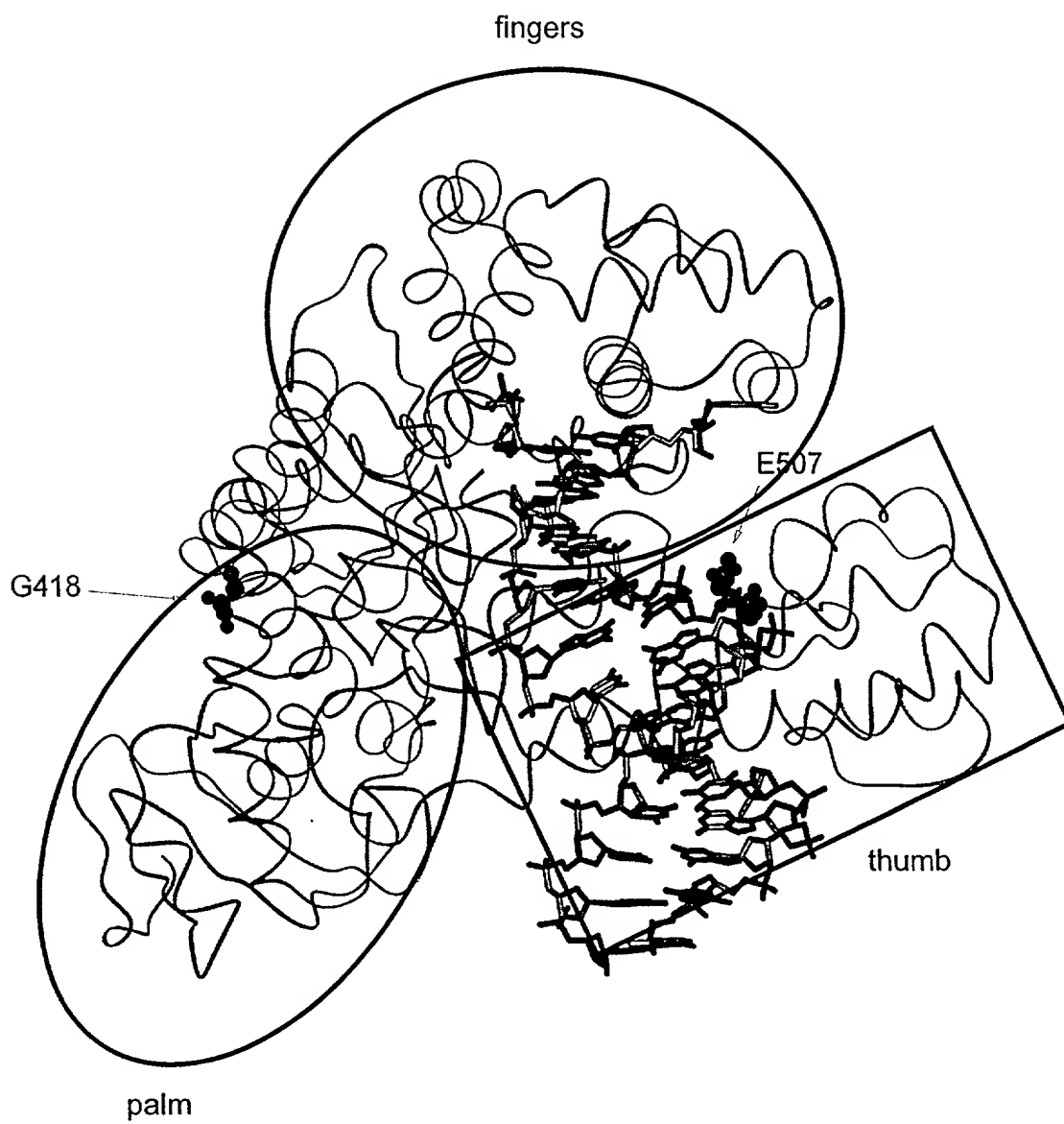


FIGURE 5

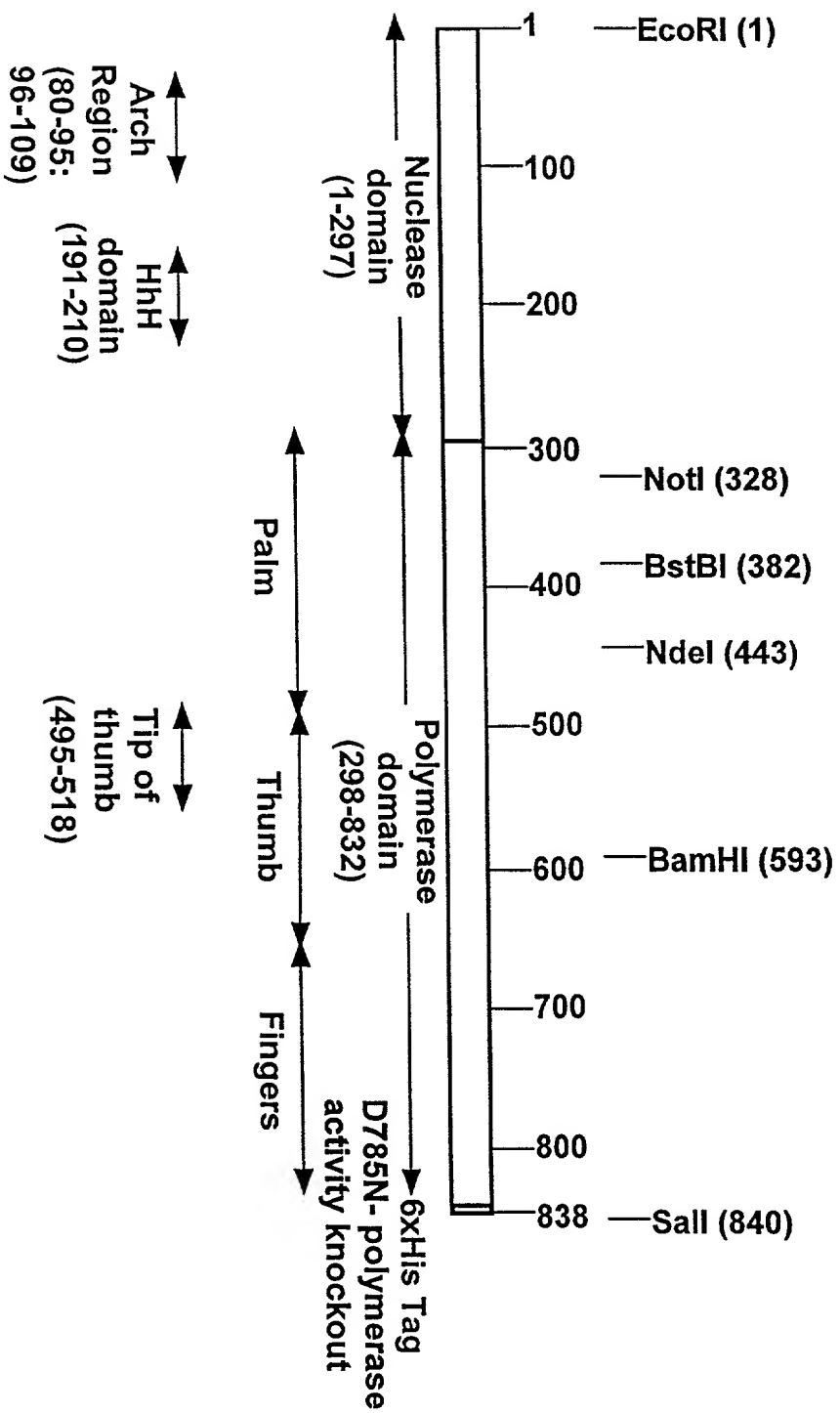


FIGURE 6

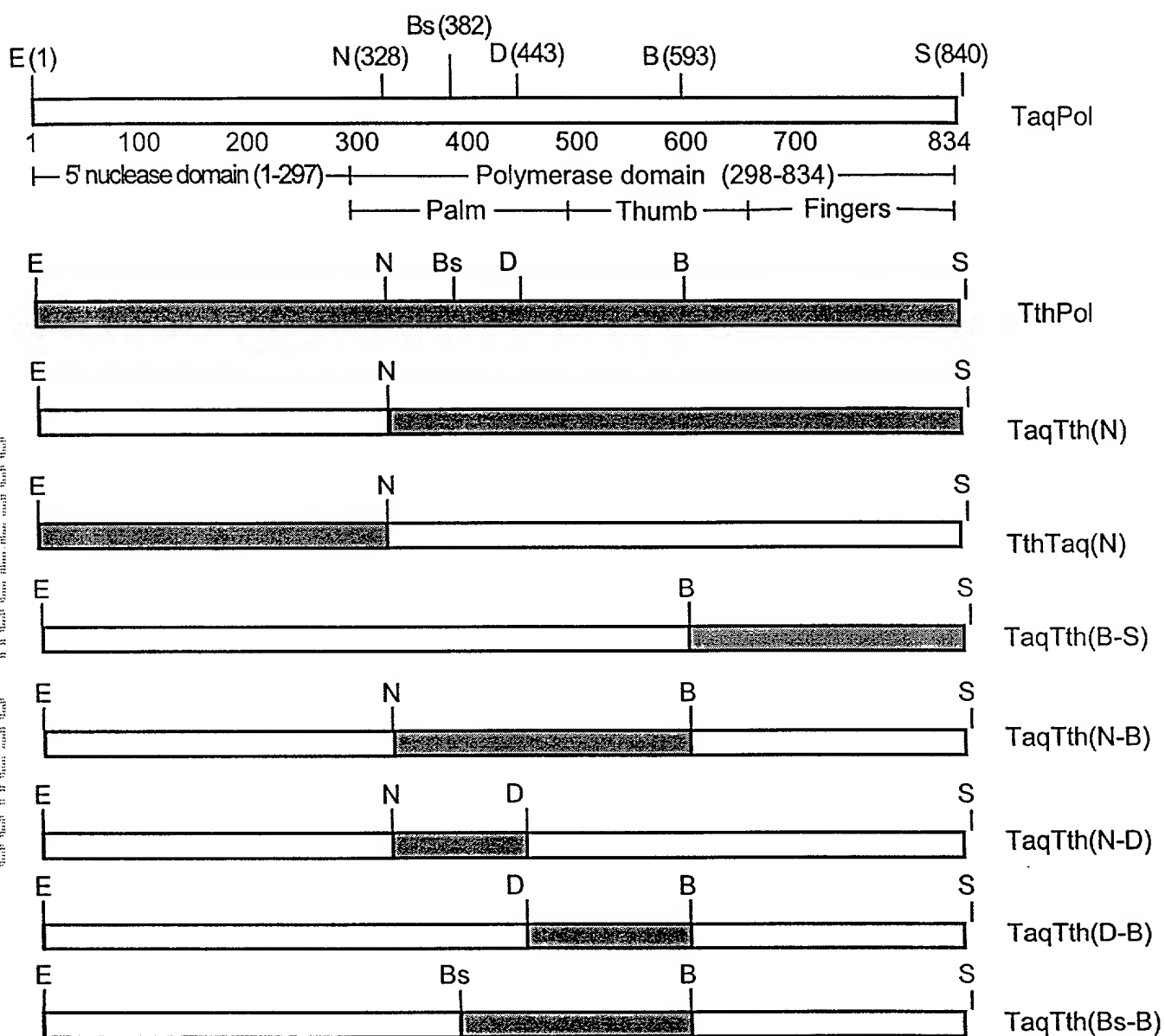


FIGURE 7

FIGURE 8A

MAJORITY [SEQ ID NO:7]	ATGXXGGGATGCTTCCGGCTCTTGAGCCCAAAGCCCGGCTCCTGCTGGTGGAGGGGACGACGCTGGCCT	
DNAPTAQ [SEQ ID NO:1]	...AG..G.....G.....G.....	70
DNAPTFL [SEQ ID NO:2]G.....G.....G.....	67
DNAPTTH [SEQ ID NO:266]	GA.....G.....A.....	70
MAJORITY	ACGGCAGCTTCTTCCGCCCTGAAGGGCTCAGCAGCCGGGGGGAACGGGTGCAGGGGGTCTACGGCTT	
DNAPTAQCA.....G.....G.....	140
DNAPTFLT.....C.....C.....T.....	137
DNAPTTHG.....G.....	140
MAJORITY	GGCCAAGAGGCTCCTCAAGGGCTGAAGGAGGAGGGGACXXGGCGGTGXTGCTGCTTTGAGGGCAAG	
DNAPTAQG.....A.....	207
DNAPTFL	A.....GT..T.....	204
DNAPTTHT..AA...C..CT.....	210
MAJORITY	GGCGCGTCTTCCGGCAGGAGGCTACGAGGCTACAAGGGGGGGGGGGGGGGGGGAGGACTTTC	
DNAPTAQG..GG.....G.....	277
DNAPTFLGA.....G.....G.....	274
DNAPTTHG.....G.....G.....	280
MAJORITY	CGCGGAGCTCGCGCTCATCAAGGAGCTGGTGGACCTCCTGGGGCTTGGGGGGCTCGAGGTCCCGGGCTA	
DNAPTAQA.....G.....G.....G.....	347
DNAPTFLG.....T.....A..C...T...G..G...T.....T.....	344
DNAPTTHT.....T..A.C.....	350

MAJORITY [SEQ ID NO:7] TCCAGGGCCACATGGAXGACCTGAXGCTCTCCTGGGAGCTXTCGAGGTGGGACCGAGCTGGCCCTGGA

DNAPTAQ	[SEQ ID NO:1]	T.....C.T....A.....G.GG.A.....	764
DNAPTFL	[SEQ ID NO:2]	GGG.....G.C...CCG..T...C.A...T.....	761
DNAPTTH	[SEQ ID NO:266]	A.....C.....C.G.....T...G...G.....C.....	770

MAJORITY GGTGGACTTCGGCAAGXGGGGGAGCGCGGACCGGGAGGGGCTTAGGGCCTTCTGGAGAGGCTGGAGTTT

DNAPTAQAA.....A.....T.....	834
DNAPTFLGG.G.C.C..CACA...A..T....T....GC...T....C.T.....	831
DNAPTTHG.....C..C.G.....D.N.A.P.T.T.H.....C.....C.....	840

MAJORITY GGCAGCGCTCCTCGACGAGTTCGGCGCTCCTGGAGGGCCCAAGCCCTCGAGGAGCGCCCGCTGGCGCGCGCG

DNAPTAQT.....AA.....	904
DNAPTFL	A.....G.....GGCA.....	T.	901
DNAPTTHG.....GGCC.....	910

MAJORITY CCGAAGGGCGCTCGTGGGCTTGTCCTTCCGGCCCGAGCCCATGTGGCCGAGCTTCTGGCCCTGGC

DNAPTAQG.....AAG.....T.....	974
DNAPTFLT..TT.....TC.T.....T.....	977
DNAPTTHG.....C.....G.....AAA.....	980

MAJORITY CCGCGCGCAGCGCGCGCGCTCCACCGCGGACGAGCCGCTTAXGGCCCTXAGGACCTXAAGGAGCTG

DNAPTAQG.....	C.C.G.T.A.AA.C..C.	C.	1044
DNAPTFL	T.GG..GT.....	G.CC..T.....C..G..G..T.	G.	1041
DNAPTTH	...TC.....C.....G.....	GGC..G.A.A..C.	C.	1050

MAJORITY [SEQ ID NO:7]	CGGGGXCTCCTCGCGAAGGACCTGGCCGTTTGGCCCTGAGGAGGGCCTXGACCTCXTGCCCGGGGACG	
DNAPTAQ [SEQ ID NO:1]G..T.....A.....AG.....C.....A.....T..G.....CC.....C.....	1114
DNAPTFL [SEQ ID NO:2]AA.....G.....G.....C.....C.....G.....T..C.....A..A.....	1111
DNAPTTH [SEQ ID NO:266]C.....C.....C.....TC.....TC.....G..A.....G.....C.....	1120
MAJORITY	ACCGCATGCTCCTCGCCATACCTCCTGGACCCCTCCAAACAGCCGCGGAGGGGTGGCCCGGGCGCTACGG	
DNAPTAQT.....T.....T.....T.....T.....T.....T.....T.....T.....	1184
DNAPTFLG.....T.....T.....T.....T.....T.....T.....T.....T.....	1181
DNAPTTHG.....T.....T.....T.....T.....T.....T.....T.....T.....	1190
MAJORITY	GGGGAGTGGACGGAGGAXCGGGGGGAGCGGGCGCTCGTCTCGGAGAGGCTCTTCXGAACCTXXGCGAG	
DNAPTAQG.....G.....G.....G.....G.....G.....G.....G.....G.....	1254
DNAPTFLT.....T.....A.....GG.....GG.....C..C.....A..C.....AAA.....	1251
DNAPTTHC.....C..C..C..C.....C..G.....CAT..G.....CGTTA..	1260
MAJORITY	CGCCTTGAGGGGAGGAGGCTCGTTTGGCTTTACGAGGAGGTGGAGAACCCCTTTCGCGGGTGGTGG	
DNAPTAQ	A..G.....G.....G.....G.....G.....G.....G.....G.....G.....	1324
DNAPTFLA.....A..A..AC..C..G.....G.....G.....G.....G.....GT.....	1321
DNAPTTHC.....A.....A.....G.....C.....A.....A.....C.....C.....	1330
MAJORITY	CGCAGATGGAGGGCAGGGGGTXXCGGCTGGACGTGGCTACCTCCAGGGCGCTXTCCTGGAGGTGGCGGA	
DNAPTAQG.....G.....G.....G.....G.....G.....G.....G.....G.....	1394
DNAPTFLGG.....C.....C.....C.....C.....C.....C.....C.....C.....	1391
DNAPTTHG.....G.....A.....A.....T.....T.....T.....T.....T.....	1400

FIGURE 8E

MAJORITY [SEQ ID NO:71]	GGAGATCGCGCGCGCTCGAGGAGGAGGTCTTCGGCGTGGCGGGGACCCCTTCAACCTCAACTCCCGGGGAG	
DNAPTAQ [SEQ ID NO:11]GC.....CC.....	1464
DNAPTFL [SEQ ID NO:21]	...G.C...AG..G.....	1461
DNAPTTH [SEQ ID NO:266]T...G.....	1470
MAJORITY	CAGCTGGAAAGGCTGCTCTTGACGAGGCTXGGGCTTCGGCGCATCGGCAAGACGGAGAGACXGGCAAGC	
DNAPTAQG.....A.....	1534
DNAPTFL	...GG.....G..C..G..T.....	1531
DNAPTTHTA.....T.G..G.....C.A.....A.....	1540
MAJORITY	GCTCGACGAGCGCGCGCTGCTGGAGCGCCCTXCGXGAGCGCGCCACCCCATCGTGGAGAAAGATCCTGCAGTA	
DNAPTAQG.....C.....C..C.....	1604
DNAPTFLT.....G..A.....CGCG.....	1601
DNAPTTHG.....A..G.....C.....C.....	1610
MAJORITY	CGGGGAGCTCAGCAAGCTCAAGAACAGCTACATXGAGCGCCCTGGCXGXCCTCGTCCACCCGAGGAGCGGGC	
DNAPTAQG...G.....T.....T.....G.A...A.....	1674
DNAPTFLA.....A.....C.C...G.....A...G.....	1671
DNAPTTHG.G.....C..AAG.....G.....	1680
MAJORITY	CGCCTGCAGACCGGCTTGAACGAGACGGCGCACGGGCGAGGCTTAGTAGCTCCGAGCGGAAAGCTGC	
DNAPTAQA.....T.....C.....	1744
DNAPTFL	..G.....C.....TCG.....	1741
DNAPTTHG.....G.....	1750

MAJORITY [SEQ ID NO:7] AGAACATCGCCGCTCGGCACCCXCTGGCCACAGGATCGCCGGGGCCTTCGTGGCCGAGGAGGGXTGGGT

DNAPTAQ [SEQ ID NO:1]G..T..G.....A..C.....G...G.. 1814
 DNAPTFL [SEQ ID NO:2]G.....T.....C..C.....A.....C.....G... 1811
 DNAPTTH [SEQ ID NO:266]CT.....G.....C...T....G 1820

MAJORITY GTTGGTGGCCCTGGACTATAGCCAGATAGAGCTCGGGGTCTGGGCCACCTCTCGGGGAGGAGAACCTG

DNAPTAQ A.....A.....G.....G.....C..... 1884
 DNAPTFL ..G.....T.....T.....T.....C..... 1881
 DNAPTTHC.....G.....C.....A..... 1890

MAJORITY ATCGGGGTCTTCAGGAGGGAGGGGACATCCACAGCCAGAGCGCCAGCTGGATGTTGGGGCTCCCCCGGG

DNAPTAQC.....GG.....G... 1954
 DNAPTFLT.....A.....T.....G.. 1951
 DNAPTTH ..A.....A.....A..... 1960

MAJORITY AGGGCGTGGACCCCTGATCGGGCGGGGGCCAGAGCATCAACTCGGGGTGCTCTAGGGCATGTCCGG

DNAPTAQG... 2024
 DNAPTFL ..A..G..A...T.....G..... 2021
 DNAPTTHGG..G.....G..... 2030

MAJORITY GCACGGCGCTCTCCGAGGAGGCTTGGCATGCCCTACGAGGAGGGGTGGCCTTCATTGAGGGCTACTCCAG

DNAPTAQA.....T.....GCA.....T... 2094
 DNAPTFLGG.....T..... 2091
 DNAPTTH ...TA..G.....T..A.....A 2100

MAJORITY [SEQ ID NO:7]	AGCTTCCGCCAAGGTGGGGCCCTGGATTGACAAGACCGCTGGAGGAGGGCAGGAGGGGGGTACGTGGAGA	2164
DNAPTAQ [SEQ ID NO:1]	2161
DNAPTFL [SEQ ID NO:2]	...A.....GG.....C.CG.....T.....	2170
DNAPTTH [SEQ ID NO:266]A.A.....G.....A.....G.....A.....	
MAJORITY	CCCTCTTGGGGCGGGGGCTACGTGCCCGACGCTCAACGGCGGGGTGAAGAGCGCTGCCGGAGGGCGGGGA	
DNAPTAQC.....A.....AG.G.....G..	2234
DNAPTFLT.....G.....C.....	2231
DNAPTTH	...AA.AA.....CA.....C.....	2240
MAJORITY	GGGCATGGCCTTCAACATGCCCGCTCCAGGGCACGGCGGGGACGTCATGAAGCTGGCCATGGTGAAGCTC	
DNAPTAQT.....	2304
DNAPTFLG.....CG...T	2301
DNAPTTHG.....	2310
MAJORITY	TTCCCGCGGGCTXCAGGAAATGGGGGCCAGGATGCTCCTXCAGGTCACGAGGAGCTGCTCCTCGAGGGCC	
DNAPTAQ	...A...GG.....T.....	2374
DNAPTFL	...T...C...G...TT.G...G.....	2371
DNAPTTH	...C.C.G...G...C.....CG...G.....	2380
MAJORITY	CGAAAGAGCGGGGAGGXGGTGGCGGCTTGGCCAAAGGAGGTCATCGAGGGGGTCTATCGCGCTGGCGGT	
DNAPTAQ	.A.....A.....CGC.....G.....	2444
DNAPTFL	...G..C.....AG...A.....GG.....CAG..	2441
DNAPTTH	..C...C.....G...A.....G.....C.....AA..C.....C.....	2450

FIGURE 8H

MAJORITY [SEQ ID NO:7] GCGCGTGGAGGTGGAGGTGGGATGGGGAGGACTGGCTCTCGGCCAAGGAGTAG

DNAPTAQ	[SEQ ID NO:1]A.....GA	2499
DNAPTFL	[SEQ ID NO:2]CG.....GT...	2496
DNAPTTH	[SEQ ID NO:266]T.....GT...	2505

TAQ PRO	S.		K.	D.	G.	PE. YKA.	A 348
TFL PRO	G. A.	L. SF.			G. WE. L.	Q. R.	G. 347
TTH PRO	A. AP.				K. C. D.	A. A. K.	350

MAJORITY [SEQ ID NO:8]		RGLLAKDLAVLALREGLDLXPGDDPMLLAYLLDPSNTTPEGVARRYGGEWTEAGEDALLSERLFXNLXX	
TAO PRO	[SEQ ID NO:4]	S.....G.P.....E.....A.....A.....WG	418
TFL PRO	[SEQ ID NO:5]	I.....F.E.....A.....QT.KE	417
TTH PRO	[SEQ ID NO:267]	S.....V.....AH.....HR..LK	420
MAJORITY		RLEGEERLLWLYXEVEKPLSRVLAHMEATGVRLDVAYLQALSLEVAEEIRRLLEEVEFRLAGHPFNLNSRD	
TAO PRO		R...R...A.....R.....A.....A.....	488
TFL PRO		K.....E.....R.....EA.V.Q.....	487
TTH PRO		K.....H.....L.....	490
MAJORITY		QLERVLFDELGLPAIGKTEKTGKRSTSAAVLEALREAHPIVEKILQYRELTCLKNTYIDPLPXLVHPRTG	
TAO PRO	S.....D.I.....	558
TFL PRO	DR.....A....K..	557
TTH PRO		R...L...Q.....H.....V....S.....	560
MAJORITY		RLHTRFNQTATGRLSSSDPNLQNI PVRTPLGQRI RRAFVAEEGWXLVALDYSQIELRVLAHLSGDENL	
TAO PRO	I.....L.....	628
TFL PRO	V..V.....	627
TTH PRO	A..A.....	630
MAJORITY		IRVFQEGRDIHTQTASWMFGVPPEAVDPLMRRAAKTINFGVLYGMSAHRLSQELAI PYEEAVAFIERYFQ	
TAO PRO		E.....R.....Q.....	698
TFL PRO		S..G.....G..S.....	697
TTH PRO		K.....V.....	700

FIGURE 9C

MAJORITY [SEQ ID NO:8] SFPKVRAWI EKTLEEGRRRGYVETLFGRRRYVPDLNARVKSUREAAERMAFNMPVOGTAAADLMK LAMVKL

TAQ PRO	[SEQ ID NO:4]E.....	768
TFL PRO	[SEQ ID NO:5]	Y.....G.....R.	767
YTH PRO	[SEQ ID NO:267]K.....	770

MAJORITY FPRLXEMGARM LQVHDEL VLEAPKXRAEXVAALAKEVMEGVYPLAVPLEVEVGXGEDWLSAKEX

TAQ PROE.....A.....R.....I.....	833
TFL PROQ.L.....D.....R.....W.Q.....L.....	831
YTH PROR.....L.....QA.....E.....A..KA.....M.....G	835

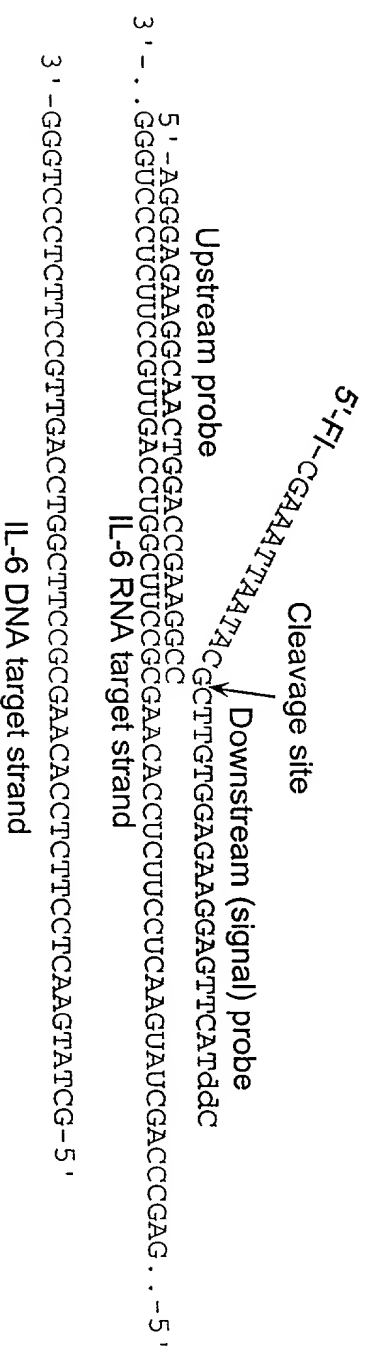
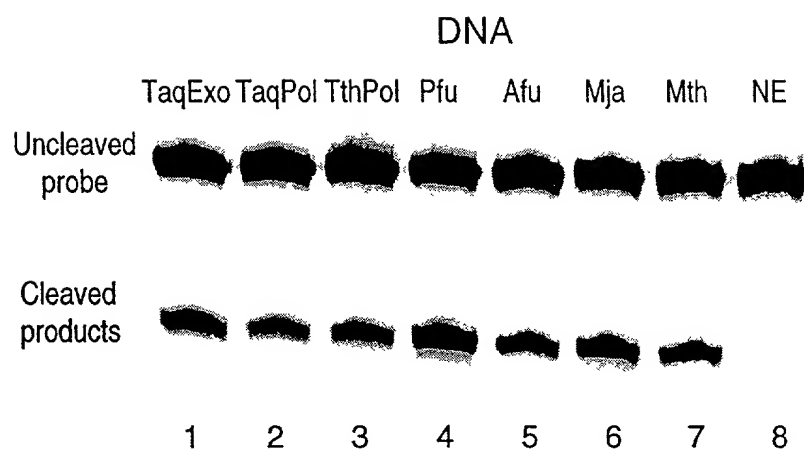


FIGURE 10

A



B

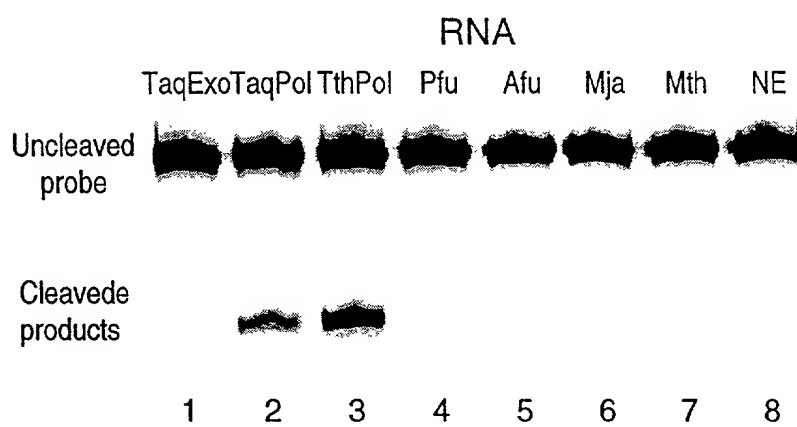


FIGURE 11

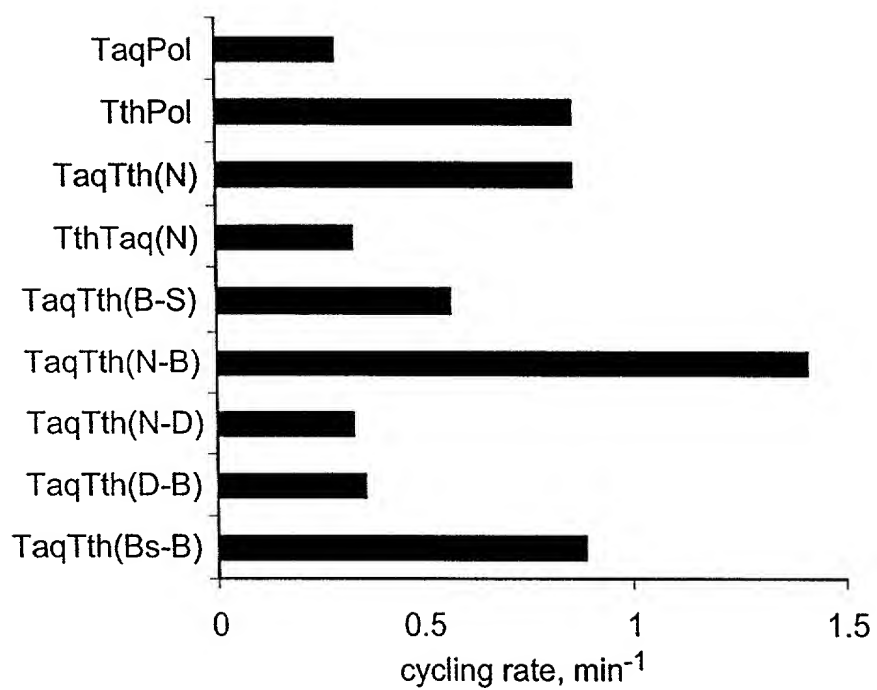


FIGURE 12

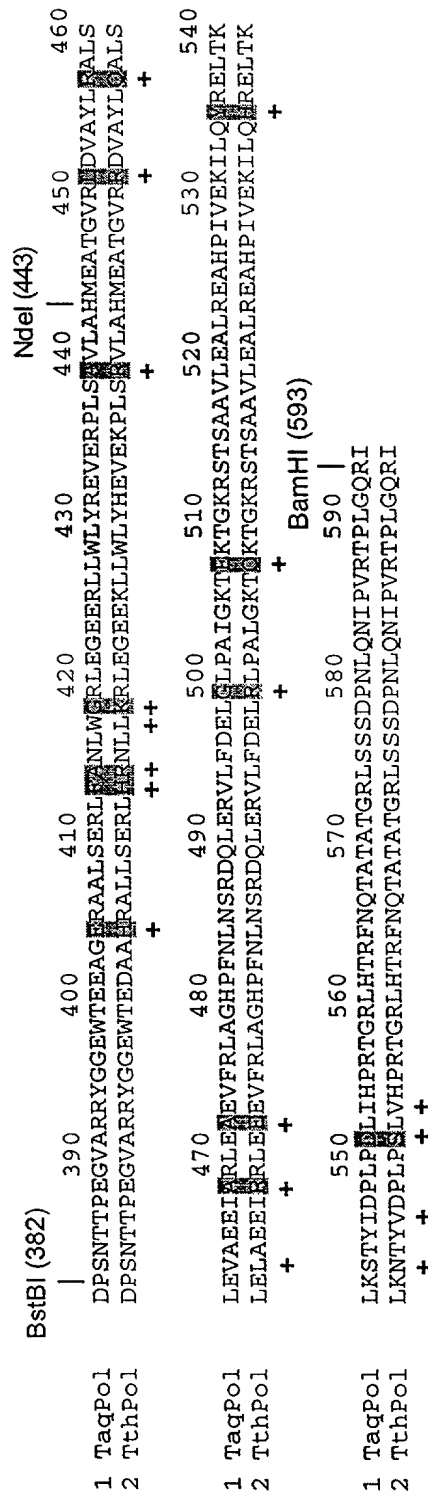


FIGURE 13

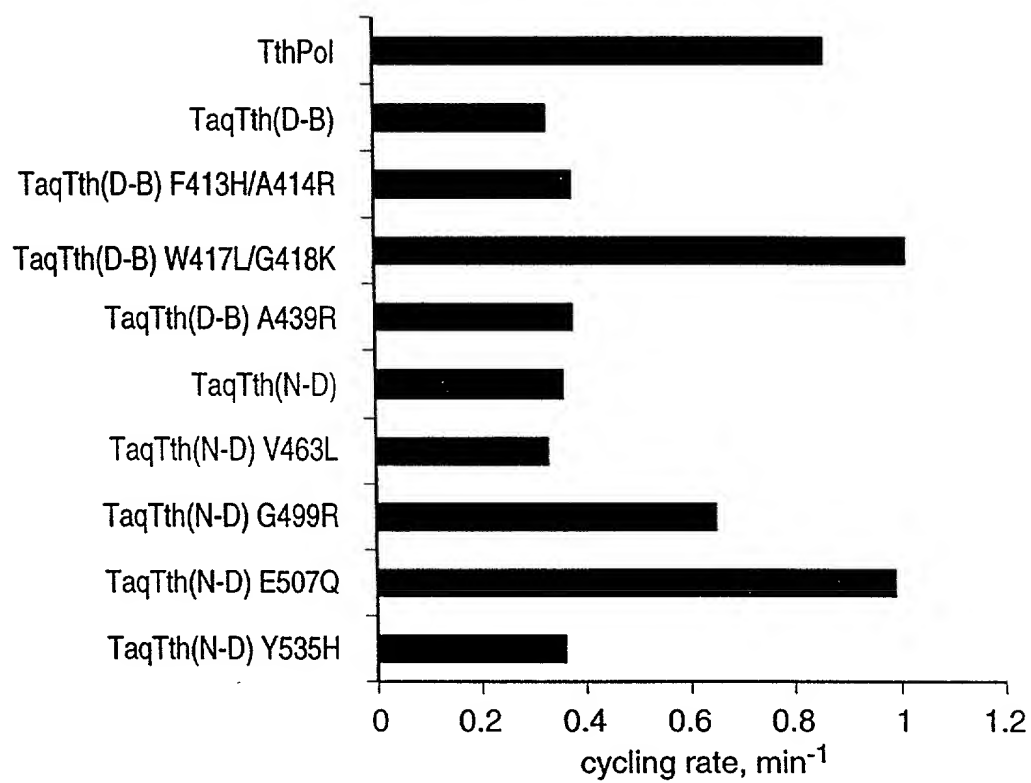


FIGURE 14

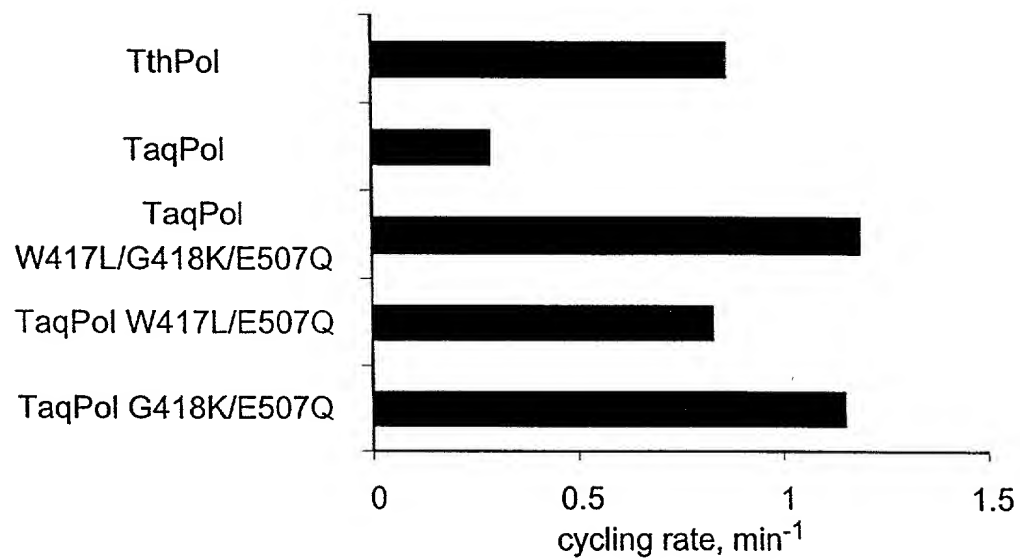


FIGURE 15


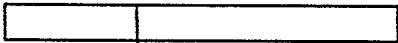




		Polymerase Activity Assays	
		<u>% Fl-labeled dUTP incorporated</u>	
		<u>RNA, p(A) or DNA, p(dA) Template</u>	
Nuclease Domain	Polymerase Domain		
		5.8 (1.00)	14.8 (1.00)
Tth			
		0.8 (0.14)	15.0 (1.01)
Taq			
		4.88 (0.84)	12.9 (0.87)
TaqTth(N)			
		0.58 (0.10)	13.3 (0.90)
TaqTth(N-B)			
		6.60 (1.14)	14.9 (1.01)
TaqTth(B-S)			
		0.42 (0.07)	12.6 (0.85)
Taq(W417L/G418K/E507Q)			

FIGURE 16



FIGURE 17

[illegible]

FIGURE 18A

CGGGAACGAGCGTCTTTG

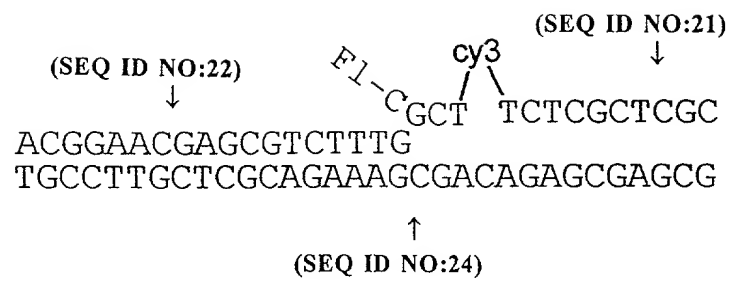


FIGURE 18B

904880"4364880

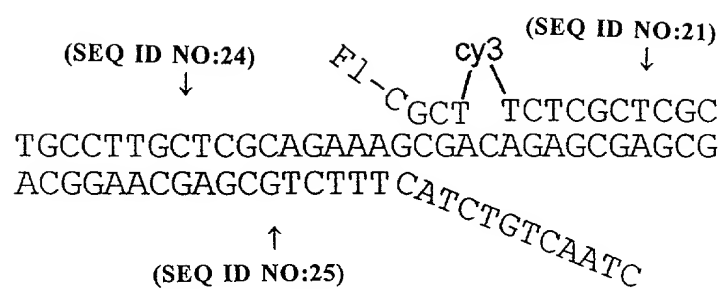


FIGURE 18C

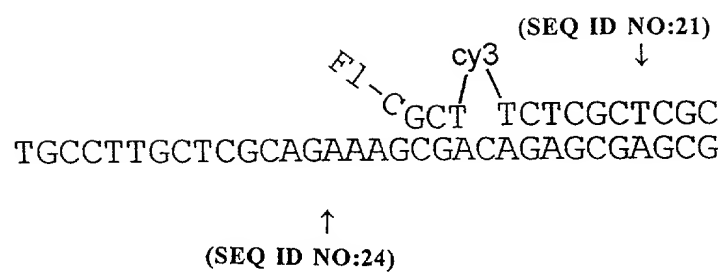


FIGURE 18D

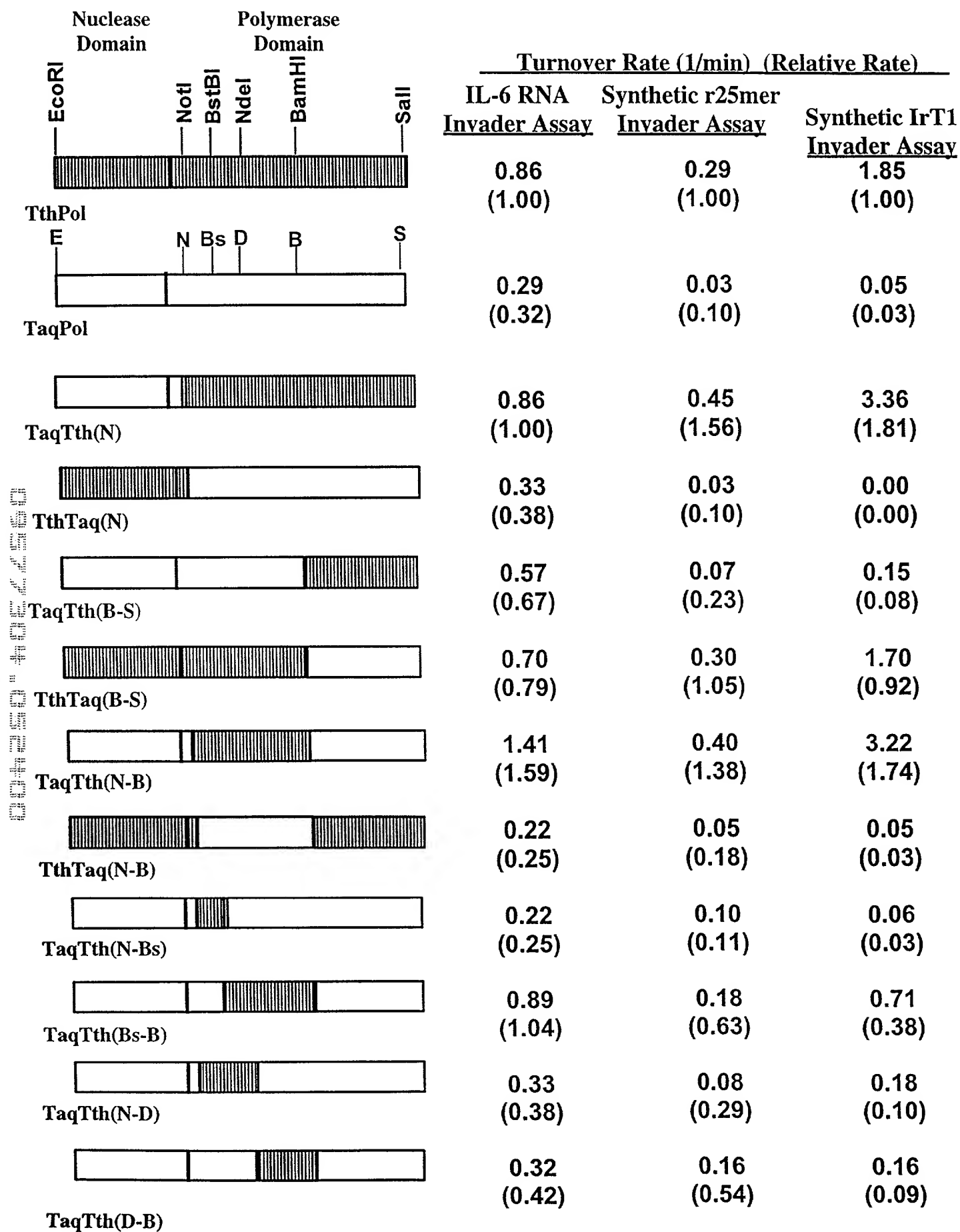


FIGURE 19

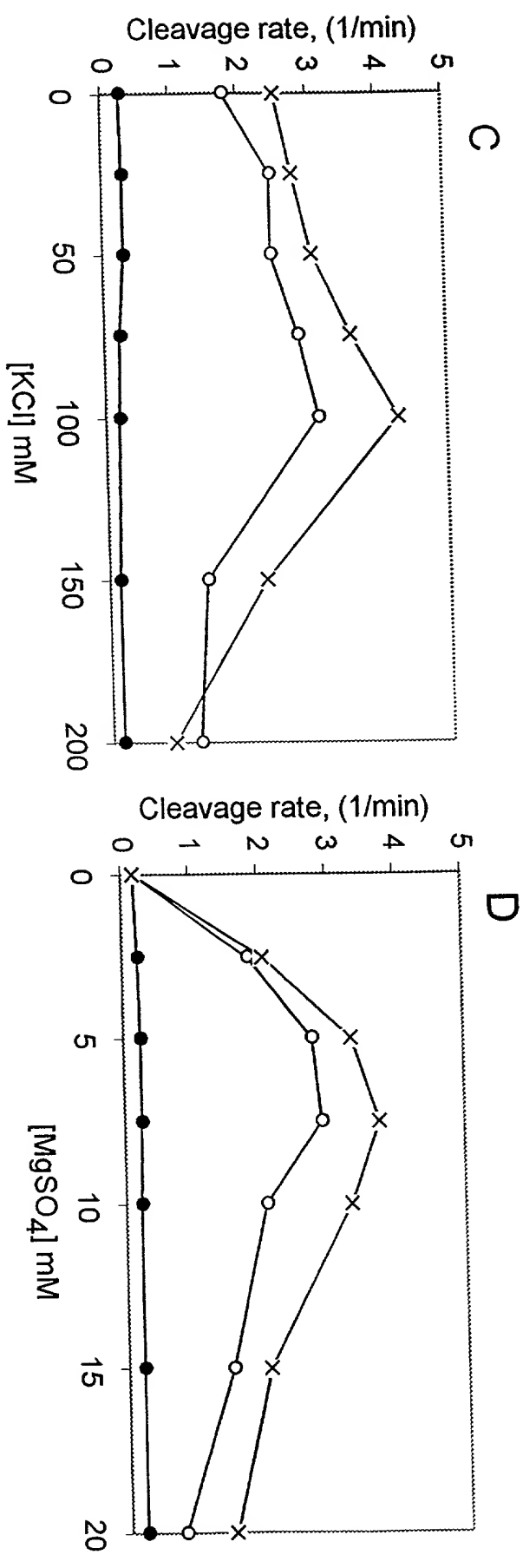
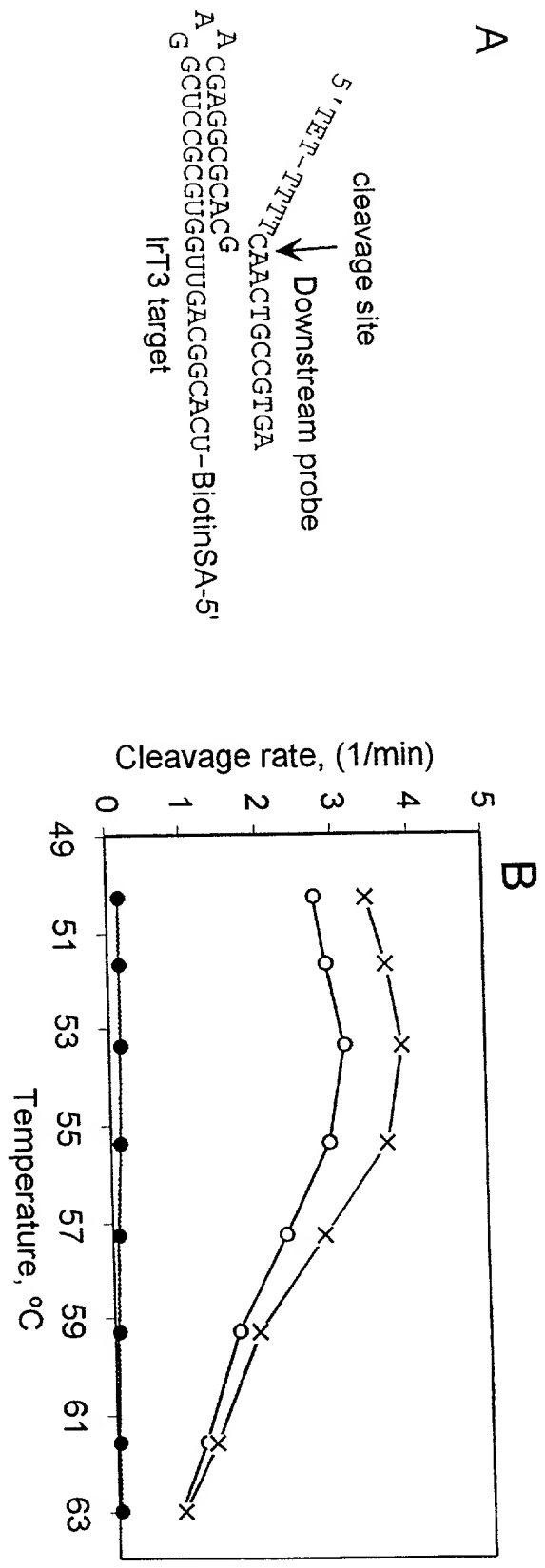


FIGURE 20

09577304.032400

FIGURE 21

A

5'-tet-TTTTCAACTGCCGTGA
A CGAGGCGCACG
A GCTCCGCGTG GTTGACGGCACT

B

5'-tet-TTTTCAACTGCCGTGA
A CGAGGCGCACG
A GCUCCGCGUGGUUGACGGCACU-BiotinSA-5'

FIGURE 22

A

(SEQ ID NO:29)



3' NH4-AATTGCTCCGCGTGGTTGACGAAGGAGGC-5'

5'-F1-TCCTTCTCAACTGCTTCCTCCG-3'



(SEQ ID NO:30)

B

(SEQ ID NO:31)



3' NH4-AATTGCTCCGCGTGGTTGACGAAGGAGGC-5'

5'-AACGAGGCGCACCTCAAATCTCCCTTT-biotin

(SEQ ID NO:29)



(SEQ ID NO:30)

00425014022500

FIGURE 23

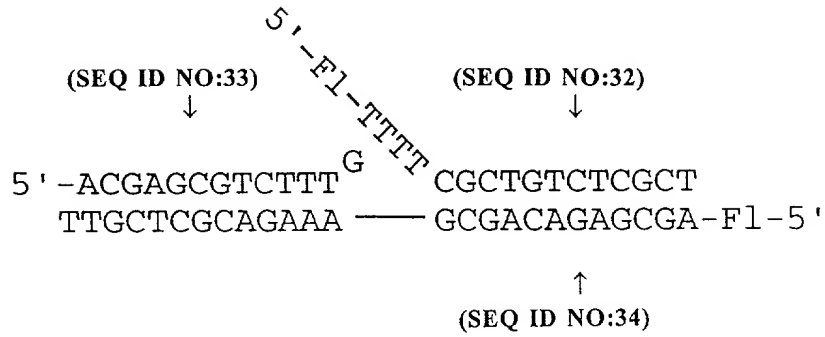
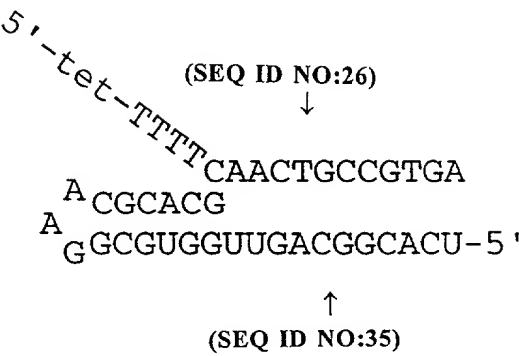
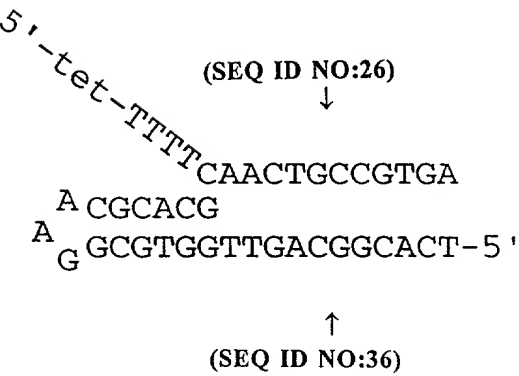


FIGURE 24

A

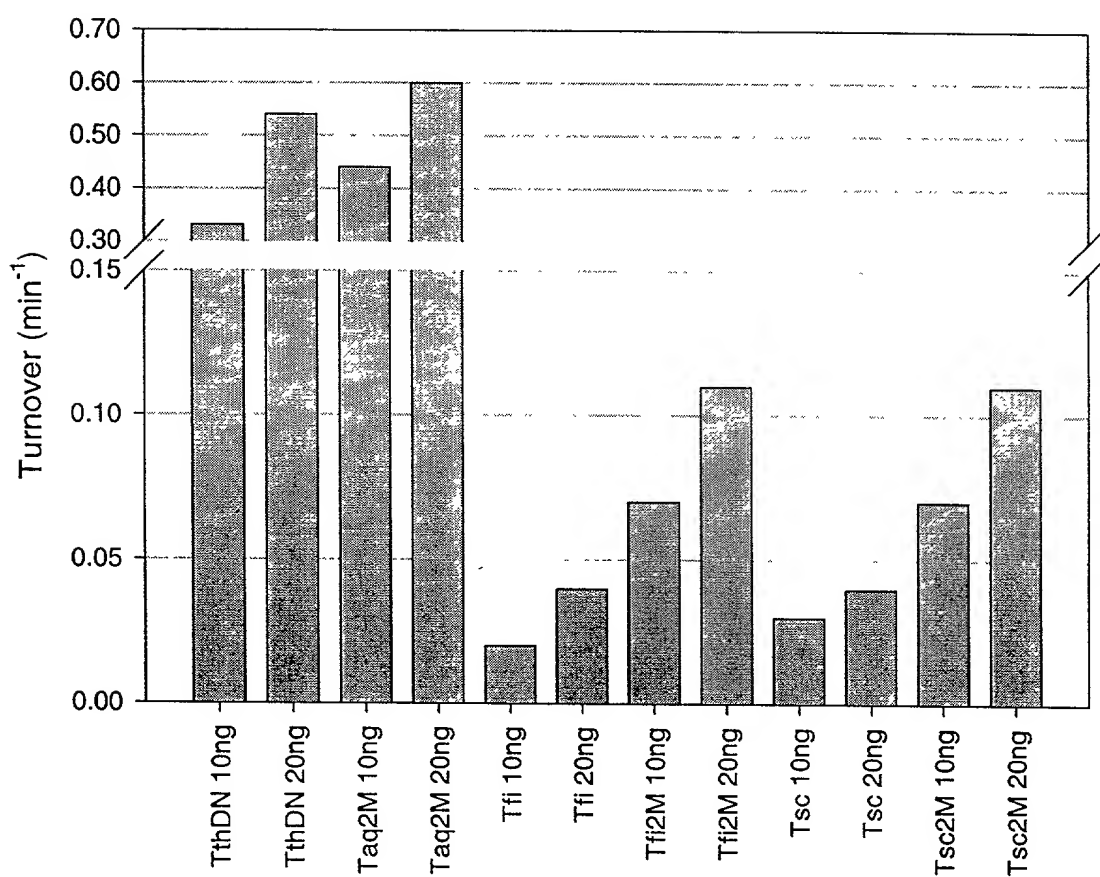


B



004200"40E2E60

FIGURE 25



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Wu-Po Ma *et al.* Group No.:
Serial No.: Examiner:
Filed:
Entitled: **Improved Enzymes For The Detection Of Specific Nucleic Acid Sequences**

POWER OF ATTORNEY BY ASSIGNEE

Assistant Commissioner for Patents
Washington, D.C. 20231

Third Wave Technologies, Inc., as Assignee of record of the entire interest of the above-identified patent application, hereby appoints the members of the firm of MEDLEN & CARROLL, LLP, a firm composed of:

Virginia S. Medlen	(Reg. No. 32,050)	J. Mitchell Jones	(Reg. No. 44,174)
Peter G. Carroll	(Reg. No. 32,837)	David J. Wilson	(Reg. No. 45,225)
Kamrin T. MacKnight	(Reg. No. 38,230)	Jason R. Bond	(Reg. No. P-45,439)
David A. Casimir	(Reg. No. 42,395)	Emily C. Toncgo	(Reg. No. P-46,473)
Maha A. Hamdan	(Reg. No. 43,655)		

as its attorneys with full power of substitution to prosecute this application and transact all business in the Patent and Trademark Office in connection therewith.

Please direct all future correspondence and telephone calls regarding this application to:

David A. Casimir
MEDLEN & CARROLL, LLP
220 Montgomery Street, Suite 2200 Telephone: 415/705-8410
San Francisco, California 94104 Facsimile: 415/397-8338

I hereby certify that the Assignment document filed with the application or filed subsequent to the filing date of the application, has been reviewed and I hereby certify that, to the best of my knowledge and belief, title is with Third Wave Technologies, Inc.

Dated: _____

By: _____

Name: _____

Title: _____

Third Wave Technologies, Inc.
502 South Rosa Road
Madison, WI 53719

DECLARATION FOR PATENT APPLICATION

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated next to my name. I believe I am an original, first and joint inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled **Improved Enzymes For The Detection Of Specific Nucleic Acid Sequences**, the specification of which is attached hereto. I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, § 1.56(a).

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment or both, under § 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of Sole or First Inventor: **Wu-Po Ma**

Inventor's Signature: _____	Date: _____
Residence: <u>1110 N. Westfield Road, Madison, WI 53717</u>	Citizenship: <u>Hong Kong</u>
Post Office Address: <u>1110 N. Westfield Road, Madison, WI 53717</u>	

Full Name of Second Joint Inventor: **Victor I. Lyamichev**

Inventor's Signature: _____	Date: _____
Residence: <u>2523 Carriedale Ct., Madison, WI 53711</u>	Citizenship: <u>Russia</u>
Post Office Address: <u>2523 Carriedale Ct., Madison, WI 53711</u>	

Full Name of Third Joint Inventor: **Michael W. Kaiser**

Inventor's Signature: _____	Date: _____
Residence: <u>2901 Shefford Dr., Madison, WI 53719</u>	Citizenship: <u>United States of America</u>
Post Office Address: <u>2901 Shefford Dr., Madison, WI 53719</u>	

DECLARATION FOR PATENT APPLICATION

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated next to my name. I believe I am an original, first and joint inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled **Improved Enzymes For The Detection Of Specific Nucleic Acid Sequences**, the specification of which is attached hereto. I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, § 1.56(a).

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment or both, under § 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of Fourth Joint Inventor: **Natalie E. Lyamicheva**

Inventor's Signature: _____ Date: _____
Residence: 2523 Carriedale Ct., Madison, WI 53711 Citizenship: Russia
Post Office Address: 2523 Carriedale Ct., Madison, WI 53711

Full Name of Fifth Joint Inventor: **Hatim Taysir Allawi**

Inventor's Signature: _____ Date: _____
Residence: 5152 Anton Dr., Apt. #315, Madison, WI 53719 Citizenship: Jordan
Post Office Address: 5152 Anton Dr., Apt. #315, Madison, WI 53719

Full Name of Sixth Joint Inventor: **James J. Schaefer**

Inventor's Signature: _____ Date: _____
Residence: 4369 Crawford Drive, Madison, WI 53711 Citizenship: United States of America
Post Office Address: 4369 Crawford Drive, Madison, WI 53711

Full Name of Seventh Joint Inventor: **Bruce P. Neri**

Inventor's Signature: _____ Date: _____
Residence: 5714 Kilkenney Place, Madison, WI 53711 Citizenship: United States of America
Post Office Address: 5714 Kilkenney Place, Madison, WI 53711